普通高等教育“十一五”国家级规划教材

卫生部“十一五”规划教材

全国高等医药教材建设研究会规划教材

全国高等学校教材

供基础、临床、预防、口腔医学类专业用

医学分子生物学

第3版

主　编　药立波

副主编　冯作化　周春燕

人民卫生出版社

全国高等学校教材

供基础、临床、预防、口腔医学类专业用

1. 医用高等数学／第5版
2. 医学物理学／第7版
3. 基础化学／第7版
4. 有机化学／第7版
5. 医学生物学／第7版
6. 系统解剖学／第7版
7. 局部解剖学／第7版
8. 组织学与胚胎学／第7版
9. 生物化学／第7版
10. 生理学／第7版
11. 医学微生物学／第7版
12. 人体寄生虫学／第7版
13. 医学免疫学／第5版
14. 病理学／第7版
15. 病理生理学／第7版
16. 药理学／第7版
17. 医学心理学／第5版
18. 法医学／第5版
19. 诊断学／第7版
20. 医学影像学／第6版
21. 内科学／第7版
22. 外科学／第7版
23. 妇产科学／第7版
24. 儿科学／第7版
25. 神经病学／第6版
26. 精神病学／第6版
27. 传染病学／第7版
28. 眼科学／第7版
29. 耳鼻咽喉-头颈外科学／第7版
30. 口腔科学／第7版
31. 皮肤性病学／第7版
32. 核医学／第7版
33. 流行病学／第7版
34. 卫生学／第7版
35. 预防医学／第5版
36. 中医学／第7版
37. 计算机应用基础／第4版
38. 体育／第4版
39. 医学细胞生物学／第4版
40. 医学分子生物学／第3版
41. 医学遗传学／第5版
42. 临床药理学／第4版
43. 医学统计学／第5版
44. 医学伦理学／第3版
45. 临床流行病学／第3版
46. 康复医学／第4版
47. 医学文献检索／第3版
48. 卫生法／第3版
49. 医学导论／第3版
50. 全科医学概论／第3版
51. 麻醉学／第2版
52. 急诊医学

策划编辑：窦天舒　祁　军
责任编辑：窦天舒　戴薇薇
封面设计：郭　鑫
版式设计：郭　鑫　李秋蓉　何美玲

人民卫生出版社 PEOPLE'S MEDICAL PUBLISHING HOUSE PMPH 遇水消失

ISBN 978-7-117-10168-4

9 787117 101684 >

定价(含光盘)：51.00元

普通高等教育“十一五”国家级规划教材
卫生部“十一五”规划教材
全国高等医药教材建设研究会规划教材

全国高等学校教材

供基础、临床、预防、口腔医学类专业用

医学分子生物学

第3版

主　编　药立波

副主编　冯作化　周春燕

编　者（以姓氏汉语拼音为序）

德　伟（南京医科大学）
冯作化（华中科技大学同济医学院）
傅　强（四川大学华西医学中心）
高国全（中山大学中山医学院）
高　旭（哈尔滨医科大学）
葛银林（青岛大学医学院）
关一夫（中国医科大学）
韩跃武（兰州大学医学院）
何凤田（第三军医大学）
焦炳华（第二军医大学）
吕建新（温州医学院）
吕社民（西安交通大学医学院）
宋方洲（重庆医科大学）
王军志（中国药品生物制品检定所）
王丽颖（吉林大学白求恩医学院）
温进坤（河北医科大学）
徐湘民（南方医科大学）
燕　秋（大连医科大学）
药立波（第四军医大学）
臧伟进（西安交通大学医学院）
张　英（复旦大学上海医学院）
赵涵芳（上海交通大学医学院）
周春燕（北京大学医学部）

人民卫生出版社

图书在版编目（CIP）数据

医学分子生物学/药立波主编. —3版. —北京：人民卫生出版社，2008.6

ISBN 978-7-117-10168-4

Ⅰ.医… Ⅱ.药… Ⅲ.医药学：分子生物学-医学院校-教材 Ⅳ.Q7

中国版本图书馆CIP数据核字（2008）第059392号

本书本印次封底贴有防伪标，请注意识别。

医学分子生物学

第3版

主　　编：药立波
出版发行：人民卫生出版社（中继线010-67616688）
地　　址：北京市丰台区方庄芳群园3区3号楼
邮　　编：100078
网　　址：http://www.pmph.com
E-mail：pmph@pmph.com
购书热线：010-67605754　010-65264830
印　　刷：潮河印业有限公司
经　　销：新华书店
开　　本：787×1092　1/16　印张：22
字　　数：590千字
版　　次：2001年9月第1版　2008年6月第3版第12次印刷
标准书号：ISBN 978-7-117-10168-4/R·10169
定价（含光盘）：51.00元

版权所有，侵权必究，打击盗版举报电话：010-87613394
（凡属印装质量问题请与本社销售部联系退换）

全国高等学校五年制临床医学专业
第七轮 规划教材修订说明

全国高等学校五年制临床医学专业卫生部规划教材从第一轮编写出版至今已有30年的历史。几十年来，在卫生部的领导和支持下，以裘法祖院士为代表的一大批有丰富临床和教学经验、有高度责任感的老教授和医学教育家参与了本套教材的创建和每一轮的修订工作，使我国的五年制临床医学教材不断丰富、完善与更新，形成了一套课程门类齐全、学科系统优化、内容衔接合理的规划教材。本套教材为推动我国医学教育事业的改革和发展做出了历史性巨大贡献。正如老一辈医学教育家亲切地称这套教材是中国医学教育的“干细胞”教材，由她衍生出了八年制和研究生两套规划教材。今天，全国一大批在临床教学、科研、医疗第一线的中青年教授、学者继承和发扬了老一辈的优良传统，积极参与了本套第七轮教材的修订和建设工作，并借鉴国内外医学教育教学的经验和成果，不断完善和提升编写的水平和质量，已逐渐将每一部教材打造成了精品，使第七轮教材更加成熟、完善和新颖。

第七轮教材的修订从2006年5月开始，其修订和编写特点如下：

●在全国广泛、深入调研基础上，总结和汲取了前六轮教材的编写经验和成果，尤其是对一些不足之处进行了大量的修改和完善，并在充分体现科学性、权威性的基础上，更考虑其全国范围的代表性和适用性。

●依然坚持教材编写“三基、五性、三特定”的原则。

●内容的深度和广度严格控制在五年制教学要求的范畴，精练文字压缩字数，以更适合广大五年制院校的要求，减轻学生的负担。

●在尽可能不增加学生负担的前提下，提高印刷装帧质量，根据学科需要，部分教材改为双色印刷、彩色印刷，以提升教材的质量和可读性。

●适应教学改革的需求，实现教材的系列化、立体化建设，本轮大部分教材配有《学习指导与习题集》、《实验指导》、《教师用书》以及配套光盘等，且与教材同期出版。

第七轮教材共52种，新增1种，即《急诊医学》。全套教材均为卫生部“十一五”规划教材，绝大部分为普通高等教育“十一五”国家级规划教材，分两批于2008年出版发行。

第七轮 教材目录

1. 医用高等数学 / 第5版　主编　张选群
2. 医学物理学 / 第7版　主编　胡新珉
3. 基础化学 / 第7版　主编　魏祖期
4. 有机化学 / 第7版　主编　吕以仙
5. 医学生物学 / 第7版　主编　傅松滨
6. 系统解剖学 / 第7版　主编　柏树令
7. 局部解剖学 / 第7版　主编　彭裕文
8. 组织学与胚胎学 / 第7版　主编　邹仲之　李继承
9. 生物化学 / 第7版　主编　查锡良
10. 生理学 / 第7版　主编　朱大年
11. 医学微生物学 / 第7版　主编　李　凡　刘晶星
12. 人体寄生虫学 / 第7版　主编　李雍龙
13. 医学免疫学 / 第5版　主编　金伯泉
14. 病理学 / 第7版　主编　李玉林
15. 病理生理学 / 第7版　主编　金惠铭　王建枝
16. 药理学 / 第7版　主编　杨宝峰
17. 医学心理学 / 第5版　主编　姚树桥　孙学礼
18. 法医学 / 第5版　主编　王保捷
19. 诊断学 / 第7版　主编　陈文彬　潘祥林
20. 医学影像学 / 第6版　主编　吴恩惠　冯敢生
21. 内科学 / 第7版　主编　陆再英　钟南山
22. 外科学 / 第7版　主编　吴在德　吴肇汉
23. 妇产科学 / 第7版　主编　乐　杰
24. 儿科学 / 第7版　主编　沈晓明　王卫平
25. 神经病学 / 第6版　主编　贾建平
26. 精神病学 / 第6版　主编　郝　伟
27. 传染病学 / 第7版　主编　杨绍基　任　红
28. 眼科学 / 第7版　主编　赵堪兴　杨培增
29. 耳鼻咽喉-头颈外科学 / 第7版　主编　田勇泉
30. 口腔科学 / 第7版　主编　张志愿
31. 皮肤性病学 / 第7版　主编　张学军
32. 核医学 / 第7版　主编　李少林　王荣福
33. 流行病学 / 第7版　主编　王建华
34. 卫生学 / 第7版　主编　仲来福
35. 预防医学 / 第5版　主编　傅　华
36. 中医学 / 第7版　主编　李家邦
37. 计算机应用基础 / 第4版　主编　邹赛德
38. 体育 / 第4版　主编　裴海泓
39. 医学细胞生物学 / 第4版　主编　陈誉华
40. 医学分子生物学 / 第3版　主编　药立波
41. 医学遗传学 / 第5版　主编　左　伋
42. 临床药理学 / 第4版　主编　李　俊
43. 医学统计学 / 第5版　主编　马斌荣
44. 医学伦理学 / 第3版　主编　丘祥兴　孙福川
45. 临床流行病学 / 第3版　主编　王家良　王滨有
46. 康复医学 / 第4版　主编　南登崑
47. 医学文献检索 / 第3版　主编　郭继军
48. 卫生法 / 第3版　主编　赵同刚
49. 医学导论 / 第3版　主编　文历阳
50. 全科医学概论 / 第3版　主编　杨秉辉
51. 麻醉学 / 第2版　主编　曾因明
52. 急诊医学　主编　沈　洪

全国高等学校临床医学专业第五届教材评审委员会

名誉主任委员　裘法祖

主任委员　陈灏珠

副主任委员　龚非力

委　　员　（以姓氏笔画为序）

于修平　王卫平　王鸿利　文继舫　朱明德　刘国良　李焕章　杨世杰

张肇达　沈　悌　吴一龙　郑树森　原　林　曾因明　樊小力

秘　　书　孙利军

第3版前言

进入21世纪以来，随着人类基因组计划的完成，医学分子生物学的发展十分迅速，使得医学院校本科生的医学分子生物学教学面临着新的挑战。在医学基础课中完成的分子生物学基础知识如何与理解临床问题相结合？在其他医学课程中学到的知识如何在分子生物学水平进一步理解？这些问题是医学分子生物学教学、也是医学基础教育中亟待解决的问题。为此我们在2004年按照全国高等医药教材建设研究会的规划，组织编写了《医学分子生物学》第2版，在当时提供了一本适合于五年制本科生教学、密切结合临床问题的医学分子生物学教学用书。本修订版一方面希望反映出医学分子生物学的迅速发展，向读者呈现该领域的新思想、新知识及其对医学的影响；另一方面希望反思过去4年间在组织教学中积累的经验和遇到的问题，在知识点的选择和介绍方面有所改进。

新版本继承了在理论问题上向临床问题倾斜，实验技术上向临床应用倾斜的基本指导思想。在编写原则上，强调在介绍基本知识、基本理论和基本技能的基础上，尽可能包含相关领域的较新进展。各章内容都尽可能补充了最新的数据和知识。例如，人类基因组结构的最新数据、微RNA在基因表达调控中的作用、细胞分化的分子机制等内容都有补充。在编写风格上，努力追求提纲挈领、条理清晰、方便教学。本版在各章开头部分先列出了关键知识点和关键词，希望能够起到导读作用。

新版仍然由三部分内容构成。第一部分是分子生物学基本原理。从基因和基因组的角度阐述基因的结构与功能、基因组DNA的复制与修复、基因的表达与调控、细胞信号转导等分子生物学的基础知识。为方便本书作为医学分子生物学的独立教学用书，第一部分的内容较上一版有所加强。“基因表达”和“基因表达调控”在新版中独立成章，其他各章内容亦作了一些基础知识的补充。第二部分则以一些危害人类健康的临床问题为线索，从分子生物学角度阐述疾病的发生和发展机制。限于篇幅，本书仅选择了部分系统的疾病和部分生理及病理过程进行介绍。选择这些问题的原则一方面是考虑到其临床重要性，另一方面是由于在分子机制方面对其已经有较多的了解，可以在一定层面和角度加以归纳和说明。本次修订中尽量避免了综述性写作，删去了部分尚未定论的内容。第三部分介绍分子生物学技术在临床上的应用，即分子诊断、分子预防和分子治疗问题。本部分在介绍基本原理和技术的同时，较系统地介绍已经在临床上使用或试用的采用分子生物学技术发展起来的诊断方法、治疗药物、治疗方案以及疫苗等。此外，鉴于疾病相关基因研究进展迅速，增加了“疾病相关基因的克隆与基因功能研究”一章。为了帮助读者对分子生物学相关药物和方案在临床治疗方面的使用和发展有更好的把握，本部分还专门安排了药物发现策略和药物研究流程及法规的内容。

本书每章末列出了部分参考书目和文献，以供读者进一步学习参考。在正文中出现的专业名词后列出了英文原名及缩写，书末列有英汉、汉英词汇表和专业名词索引，便于读者查阅。需要指出的是，由于分子生物学专业新词汇层出不穷，其中有些词汇各家译法不一，尚无公认的译名，只能以英文原名列出，或使用暂译名。

本书读者对象定位于高等医药院校五年制本科生及临床医生，主要作为医学本科生的医学分子生物学必修课或选修课教材，同时可以作为临床医生扩展医学分子生物学知识的参考用书，以及国家执业医师考试和硕士研究生入学考试的医学分子生物学参考用书。

本教材由来自全国 22 所高等院校和科研院所的 23 位教授分工编写。作者主要为在生物化学与分子生物学教学和科研第一线工作的人员，所承担的章节大部分都是作者所直接从事的研究领域；有关药物发展和相关法规的编写则由药学专家和国家药品检定专家分别负责。

本教材在编写工作中得到了全国高等院校教材建设研究会、国家卫生部教材办公室、第四军医大学训练部教务处、第四军医大学基础部、北京大学基础医学院、华中科技大学同济医学院的大力支持。第四军医大学生物化学与分子生物学教研室秘书王景媛女士在插图和文字整理方面做了大量工作，在此一并致谢。

药立波　冯作化　周春燕

2008 年 3 月于西安

目　录

绪论 …… 1

一、医学分子生物学定义 / 1

二、医学分子生物学发展过程中的重要事件 / 1

三、医学分子生物学的现状与未来 / 3

第一篇　分子生物学基本原理

第一章　基因的结构与功能 …… 6

第一节　基因的化学结构 / 6

一、DNA 和 RNA 的化学组成 / 7

二、DNA 的一级结构意指分子中的核苷酸排列顺序 / 7

三、双链反向互补的右手螺旋是 DNA 的主要二级结构 / 8

四、DNA 在细胞内组装为致密结构 / 9

第二节　基因的功能和组构 / 12

一、结构基因为多肽链和特定 RNA 分子编码 / 12

二、基因包含结构基因和调控序列 / 13

第三节　RNA 的结构与功能 / 16

一、信使 RNA 具有特征结构 / 16

二、转运 RNA 有氨基酸接纳茎和反密码环 / 17

三、核糖体 RNA 是核糖体的组成成分 / 18

四、小分子 RNA 有不同的种类和功能 / 19

第四节　核酸酶 / 19

第二章　基因组的结构与功能 …… 20

第一节　人基因组的结构与功能 / 21

一、人基因组中只有不到三万种基因 / 21

二、人基因组中存在大量重复序列 / 22

三、多基因家族与假基因 / 24

四、线粒体 DNA 结构有别于染色体 DNA / 24

第二节　原核生物基因组的结构与功能 / 25

一、原核生物基因以操纵子方式组构 / 25

二、原核生物中的质粒 DNA / 25

第三节　病毒基因组的结构与功能 / 26

第三章　基因组复制 …… 28

第一节　基因组 DNA 复制的基本过程和共同特点 / 28

一、DNA 复制是酶促化学反应 / 29

二、DNA 复制的基本过程 / 29

三、DNA 复制过程的共同特点 / 30
四、DNA 复制的保真性 / 31
第二节 原核生物 DNA 的复制 / 31
一、大肠杆菌 DNA 聚合酶 / 32
二、大肠杆菌基因组 DNA 复制的基本过程 / 32
三、原核生物 DNA 复制的调节 / 34
第三节 真核生物染色体 DNA 的复制 / 35
一、真核生物染色体 DNA 与原核生物 DNA 复制的主要不同点 / 35
二、真核生物染色体 DNA 复制过程 / 35
三、真核生物 DNA 复制的调控 / 39
四、线粒体 DNA 的复制 / 39
第四节 某些病毒基因组的复制 / 41
一、腺病毒 DNA 的复制 / 41
二、单链 RNA 病毒基因组的复制 / 41
三、反转录病毒基因组的复制 / 42
四、乙型肝炎病毒基因组的复制 / 42

第四章 DNA 损伤与修复 …… 43
第一节 DNA 损伤 / 43
一、导致 DNA 损伤的多种因素及其机制 / 44
二、DNA 损伤以多种类型存在 / 46
第二节 DNA 损伤的修复 / 47
一、某些 DNA 损伤可以直接修复 / 47
二、切除修复是最普遍的 DNA 损伤修复方式 / 49
三、DNA 严重损伤时需要进行重组修复 / 51
四、跨越 DNA 损伤的复制后修复 / 52
第三节 DNA 损伤和修复的意义 / 54
一、DNA 损伤具有双重效应 / 54
二、DNA 损伤修复障碍与多种疾病相关 / 54

第五章 基因表达 …… 56
第一节 RNA 的转录合成 / 56
一、基因组 DNA 是 RNA 合成的模板 / 57
二、DNA 指导的 RNA 聚合酶催化 RNA 合成 / 57
三、转录起始复合体在启动子部位形成 / 58
四、RNA 链随着转录泡的移动得以延长 / 60
五、转录终止受 DNA 模板上终止信号的控制 / 60
六、初级转录物经转录后加工方可成熟 / 61
第二节 蛋白质的生物合成——翻译 / 64
一、蛋白质合成体系 / 65
二、氨基酰 tRNA 合成酶保证氨基酸与 tRNA 的正确结合 / 66
三、起始复合物的装配启动新生肽链合成 / 67
四、新生肽链在核糖体上以连续循环形式进行延伸 / 69

五、终止密码子导致肽链合成停止 / 70
六、新生肽链经加工后成为具有功能的蛋白质 / 71
七、真核生物蛋白质合成的其他特点 / 72

第六章 基因表达调控 …… 74
第一节 基因表达调控的基本规律 / 74
第二节 原核生物的基因表达调控 / 76
一、转录水平的调控是原核生物基因表达的关键环节 / 76
二、翻译水平的调控是对转录调控的补充 / 79
第三节 真核生物的基因表达调控 / 81
一、真核生物染色质结构直接影响基因转录 / 81
二、转录速度决定 RNA 合成效率 / 82
三、真核生物在转录后仍可控制 mRNA 的结构和功能 / 85
四、翻译水平的调控决定蛋白质合成速度 / 87
第四节 基因表达调控异常与疾病 / 87

第七章 基因组学及相关组学 …… 88
第一节 基因组学 / 89
一、基因组学的概念 / 89
二、人类基因组计划 / 89
第二节 后基因组学 / 92
一、功能基因组学 / 92
二、转录组学 / 93
第三节 蛋白质组学 / 93
一、蛋白质组和蛋白质组学的概念 / 93
二、蛋白质组与基因组的比较 / 94
三、蛋白质组学研究的内容和方法 / 95
四、蛋白质组数据库 / 95
五、蛋白质组学的应用及进展 / 96
第四节 后基因组时代生命科学的发展趋势 / 96
一、基因组学研究的深入 / 96
二、基因组学相关领域 / 97
三、基因组与后基因组研究将引发医学革命 / 98

第八章 细胞信号转导 …… 99
第一节 细胞信号转导的分子机制 / 99
一、细胞信号转导需要多种分子的相互作用 / 99
二、信号转导分子以两种基本方式向下游传递信号 / 101
第二节 不同受体介导的信号转导途径 / 103
一、细胞内受体直接传递信号 / 103
二、离子通道型受体通过控制离子流动而传递信号 / 104
三、G 蛋白和小分子信使是 G 蛋白偶联受体介导的信号转导

途径的主要特征 / 104
四、蛋白激酶偶联受体介导的信号转导途径主要涉及蛋白质的相互作用 / 109
第三节 信号转导的基本规律和复杂性 / 111
一、各种信号转导机制具有共同的基本规律 / 111
二、一条信号转导途径中的功能分子可影响和调节其他途径 / 111
三、细胞信号转导复杂且具有多样性 / 112

第九章 细胞增殖分化的分子机制 …… 114
第一节 细胞增殖的分子基础 / 114
一、细胞增殖是生长的主要因素之一 / 114
二、细胞周期是细胞分裂过程中的有序事件 / 115
三、细胞周期蛋白调节细胞周期中不同时相的正确转换 / 115
四、不同时相转换的关卡是细胞周期调控的关键环节 / 116
五、生长因子具有调节细胞生长与分化的作用 / 118
第二节 细胞分化的分子机制 / 120
一、细胞分化是一渐进过程并受到一系列细胞信号分子的调控 / 120
二、胚胎干细胞分化具有全能性 / 123
三、成体细胞的横向分化 / 125
四、再生医学是21世纪医学研究的热点 / 126

第二篇 疾病的分子生物学基础

第十章 基因变异与疾病 …… 127
第一节 基因变异的基本概念 / 128
一、基因变异包括DNA多态性和致病突变 / 128
二、人基因组中有多种类型的基因变异 / 129
三、基因变异是可以遗传的 / 130
第二节 基因变异的生物学效应 / 131
一、DNA变异可使基因功能减弱 / 131
二、DNA变异可使基因功能增强 / 135

第十一章 细胞异常增生性疾病的分子机制 …… 140
第一节 肿瘤细胞增殖机制 / 140
一、癌基因 / 141
二、抑癌基因 / 143
三、癌基因与抑癌基因在肿瘤发生中的作用 / 146
第二节 器官纤维化的分子机制 / 147

第十二章 感染性疾病的分子生物学 …… 149
第一节 病原菌致病的分子机制 / 149
一、病原菌的基因组 / 149

二、病原菌毒素及其编码基因 / 150
三、细菌毒素以各种方式影响宿主细胞功能 / 151
四、抗药性基因的出现使细菌产生耐药 / 152
第二节 病毒致病的分子机制 / 154
一、病毒通过病毒受体进入宿主细胞 / 154
二、病毒感染宿主的基本过程 / 155
三、病毒感染对宿主细胞功能的影响 / 156
第三节 典型感染性疾病的分子机制 / 158

第十三章 炎症的分子机制 …… 161
第一节 炎症刺激信号及其受体 / 161
一、生物性致炎因素的主要受体——TLR 家族 / 161
二、理化伤害性刺激信号受体 / 162
第二节 参与炎症反应的细胞及其作用机制 / 163
一、参与炎症反应的细胞 / 163
二、炎症反应细胞聚集的分子机制 / 163
第三节 主要炎症介质及其作用机制 / 166
一、细胞释放的炎症介质 / 166
二、体液中的炎症介质 / 169
第四节 炎症反应相关的信号转导机制 / 170
第五节 炎症反应的调控点 / 172
一、早期炎症正向信号 / 172
二、早期炎症终止信号 / 173
三、组织损伤转向组织修复的信号 / 174
四、中枢神经系统在炎症反应中的作用 / 174

第十四章 心血管系统疾病的分子机制 …… 175
第一节 心血管活性物质与高血压 / 175
一、心血管活性肽分为缩血管活性肽和舒血管活性肽 / 175
二、缩血管活性肽和舒血管活性肽之间的平衡失调引起高血压 / 175
三、一些气体分子是重要的心血管活性物质 / 177
第二节 动脉粥样硬化的发病机制与分子基础 / 177
一、血管平滑肌细胞转化和增殖的细胞外信号 / 178
二、控制血管平滑肌细胞表型的转录因子 / 179
三、高脂血症是动脉粥样硬化发生发展的重要原因 / 180
四、慢性炎症反应贯穿于动脉粥样硬化发生发展的全过程 / 180
五、活性氧促进动脉粥样硬化的发生与发展 / 181
第三节 心肌肥厚的分子机制 / 182
一、心血管活性肽和生长因子引起心肌肥厚 / 182
二、心血管活性肽和生长因子致心肌肥厚效应的信号转导途径 / 182
第四节 心律失常的分子机制 / 183
一、编码心肌离子通道蛋白的基因突变引起 LQTS / 183
二、多种因素参与房颤的发生 / 184

第十五章 内分泌及代谢疾病的分子机制 …… 186
第一节 内分泌代谢系统的分子生物学特点 / 186
一、内分泌系统的进化 / 186
二、内分泌系统发育的分子调控 / 187
三、肽类激素合成与分泌的分子调控特点 / 189
四、内分泌代谢疾病的分子生物学机制 / 190
第二节 内分泌代谢疾病的分子机制举例 / 191
一、侏儒症 / 191
二、糖尿病的分子机制 / 192

第十六章 免疫系统疾病的分子机制 …… 195
第一节 免疫缺陷病的分子机制 / 195
一、抗体基因的结构与表达异常引起 B 细胞免疫缺陷病 / 196
二、TCR 基因结构与表达异常引起 T 细胞免疫缺陷病 / 199
三、T 细胞和 B 细胞异常引起严重联合免疫缺陷病 / 200
第二节 自身免疫病的分子机制 / 201
一、HLA 的基因组成及表达特点决定其功能 / 201
二、HLA 多态性与自身免疫病的易感性密切相关 / 202

第十七章 应激反应的分子机制 …… 206
第一节 概述 / 206
一、应激原是诱发应激反应的因素 / 206
二、全身应激反应是多种信号分子协调体内多种细胞参与的全身反应 / 206
三、一系列高度有序的细胞内事件构成细胞应激反应 / 207
第二节 全身应激反应的相关分子 / 207
一、激素是全身应激反应时产生的细胞间通讯信号分子 / 207
二、急性期反应蛋白是在非特异性应激反应中大量产生的防御性蛋白 / 208
三、热休克蛋白是在应激状态下保护细胞的蛋白成分 / 209
第三节 细胞应激反应的分子机制 / 211
一、细胞内 DNA 损伤诱发细胞应激反应 / 211
二、低氧可诱发细胞应激反应 / 213
三、活性氧和自由基可诱发细胞应激反应 / 214

第十八章 衰老和老年退行性疾病的分子机制 …… 216
第一节 概述 / 216
一、衰老是基因突变积累导致的细胞、组织和器官功能衰退 / 216
二、分子改变是衰老的基础 / 217
第二节 衰老过程受特定的基因控制 / 217
一、衰老相关基因影响机体衰老的进程 / 218
二、生物体存在延长寿命的长寿相关基因 / 218

第三节 线粒体 DNA 与衰老 / 219
一、线粒体 DNA 异常是线粒体病的分子基础 / 220
二、线粒体 DNA 突变是衰老自由基学说的核心 / 220
第四节 老年退行性疾病的分子机制 / 220
一、早老性痴呆 / 221
二、帕金森病 / 223

第三篇 分子诊断、预防和治疗

第十九章 基因操作 …… 225
第一节 核酸的理化性质 / 226
一、核酸的一般理化特性 / 226
二、核酸分子杂交的原理是 DNA 的变性与复性 / 227
第二节 基因的获取 / 228
一、用 PCR 技术获取基因 / 228
二、用反转录-PCR 技术获取基因 / 231
三、从基因文库中获取基因 / 231
四、人工合成基因 / 233
第三节 基因的克隆 / 233
第四节 克隆基因的表达 / 235
一、原核表达系统 / 235
二、真核表达系统 / 235
第五节 基因结构的分析及改造 / 236
一、基因的一级结构分析 / 236
二、基因的改造 / 238
第六节 基因的拷贝数分析 / 239
第七节 基因表达的分析 / 240
一、Northern 印迹定性分析基因的表达 / 240
二、RT-PCR 定量分析基因的表达 / 240
三、实时 RT-PCR 定量分析基因的表达 / 240
第八节 基因功能的分析 / 241
一、转基因技术用于基因功能的研究 / 242
二、基因打靶技术用于基因功能的研究 / 243
三、RNA 干扰技术用于基因功能的研究 / 244

第二十章 疾病相关基因的克隆与基因功能研究 …… 246
第一节 鉴定克隆疾病相关基因的原则 / 246
一、鉴定克隆疾病相关基因的关键是确定疾病表型和基因间的实质联系 / 246
二、鉴定克隆疾病相关基因需要多学科多途径的综合策略 / 247
三、确定候选基因是多种克隆疾病相关基因方法的交汇 / 247
第二节 基因相关疾病克隆的策略和方法 / 248
一、不依赖染色体定位的疾病相关基因克隆策略 / 248

二、定位克隆是后基因组时代的常用策略 / 251
三、利用生物信息数据库有助于疾病相关基因克隆的研究 / 253
第三节 疾病相关基因的功能研究 / 253
一、基因比对及功能诠释 / 254
二、利用工程细胞研究基因功能 / 254
三、研究蛋白质分子间的相互作用，通过了解生物学途径确定基因功能 / 254
四、基因修饰动物是整体研究基因功能的重要方法 / 255

第二十一章 基因诊断 …… 256
第一节 基因诊断的概念 / 257
一、DNA 是基因诊断的主要对象 / 257
二、基因诊断包括定性和定量分析 / 257
三、基因诊断操作的四个基本步骤 / 257
第二节 基因诊断技术的方法学 / 258
一、利用遗传标志连锁分析间接检测遗传病 / 258
二、针对遗传病致病突变的直接诊断 / 261
第三节 基因诊断的医学应用 / 265
一、基因检测是疾病诊断和风险预测的有力工具 / 265
二、基因诊断是遗传病预防的重要技术手段 / 267
三、基因诊断可用于疗效评价和用药指导 / 268
四、DNA 指纹鉴定是法医学个体识别的核心技术 / 268

第二十二章 基因工程药物与疫苗 …… 270
第一节 基因工程药物和疫苗的种类及制备体系 / 270
一、基因工程药物与疫苗的种类 / 270
二、基因工程药物和疫苗的制备体系 / 270
第二节 基因工程蛋白质/多肽药物 / 271
一、细胞因子类药物 / 271
二、激素类药物 / 272
三、基因工程抗体药物 / 273
四、溶栓和抗凝血药物 / 274
五、其他活性蛋白类药物 / 275
第三节 基因药物 / 277
第四节 基因工程疫苗 / 279
一、病毒疫苗 / 279
二、细菌疫苗 / 280
三、寄生虫疫苗 / 281
四、肿瘤疫苗 / 281
第五节 基因工程药物与疫苗的发展前景 / 282

第二十三章 基因治疗 …… 284

第一节 基因治疗的策略和基本程序 / 285
一、基因治疗可修复错误基因或干预异常表达的基因 / 285
二、利用载体将目的基因导入到受体细胞表达是基因治疗的基本程序 / 285
三、病毒载体和非病毒载体均可用于导入目的基因 / 286
四、基因干预可以抑制特定基因的表达 / 292
第二节 基因治疗的临床应用现状 / 292
一、用正常基因替代错误基因治疗遗传病 / 292
二、肿瘤的基因治疗可在多个环节进行 / 292
三、心脑血管病的基因治疗 / 294
四、HIV 感染的基因治疗策略是抑制病毒复制或提高机体的免疫功能 / 294
五、开展基因治疗必须遵守国家法规 / 295

第二十四章 分子生物学与药物发现 …… 296
第一节 药物发现模式的变革 / 296
一、传统的药物发现模式 / 297
二、现代药物发现策略 / 297
三、未来药物发现模式 / 298
第二节 分子生物学技术在药物发现中的应用 / 299
一、药物靶点 / 299
二、分子生物学研究提供新的药物靶点 / 299
三、分子生物学技术在药物靶点的确认与验证中的应用 / 300
四、分子生物学技术在药物筛选中的应用 / 301
第三节 现代生物学与药物发现 / 302
一、药物基因组学 / 302
二、药物蛋白质组学 / 303
三、其他现代生物学方法与药物发现 / 303

第二十五章 医药生物技术产品的研究开发与管理 …… 306
第一节 医药生物技术产品研究开发的一般过程 / 306
一、实验室研究 / 306
二、中试研究 / 306
三、临床前研究 / 307
四、临床研究 / 308
五、新药临床研究与生产的申报和审批 / 309
第二节 医药生物技术产品研究开发的相关法规 / 310
一、医药生物技术产品研究开发相关法规的主要内容 / 310
二、药物非临床研究质量管理规范 / 312
三、药物临床试验质量管理规范 / 313
四、药品生产质量管理规范 / 315
第三节 医药生物技术及其产品的生物安全问题 / 316
一、医药生物技术及其产品生物安全问题的范围 / 316

二、医药生物技术及其产品生物安全问题的预防和控制 / 317

参考书目 …… 319

中文索引 …… 321

英文索引 …… 328

绪　论

疾病始终与人类的进化和发展并存，自然在赋予人类各种生存优势的同时，也不断带来创伤、疾病和死亡的痛苦。古往今来，人类所从事的科学研究一方面是发明各种工具和用品以改善生存条件，另一方面就是设法减少疾病带来的痛苦和死亡以提高生存质量和时间，后者就是医学。医学在人类自身及社会发展中的价值是无法估量的。

医学科学的发展，是人类同疾病及影响健康的一切不利因素进行不间断斗争的经验总结和循序提高的过程，是不断拓展新的探索领域以求不断进步的过程。从以经验和哲学思考为主导的几大古代医学体系的建立，到以实验、解析、认知为主导的近代医学实践，医学已经走过了从混沌茫然到有所把握、从被动感知到主动探索、从经验主体到技术主体的阶段。如今，主导 21 世纪生命科学前沿的分子生物学的发展已经引领现代医学进入了分子医学时代。

一、医学分子生物学定义

医学分子生物学是分子生物学的重要分支。她在医学领域内探讨分子生物学的基本理论，发展分子生物学的基本技术；致力于阐明生物大分子结构、功能、调控机制以及人体各种生理和病理状态的分子机制，她的发展无疑将推动新的诊断、治疗和预防方法以及新的健康理念的建立。

单纯地认识各种生物分子（蛋白质、核酸等）的特性，并不能代表分子生物学的全部含义和任务。把握所有这些分子的特性和它们的联系，从分子角度解释生命科学的最基本问题，诸如生命的稳态、生命的存活与死亡、生命的繁殖、生命的发生及进化的机制才是分子生物学的真正价值体现。阐明生命物质，尤其是生物信息大分子的结构、功能及它们所构成的信息系统的流动和整合如何形成了生命——自然界这种特殊的物质形式，将是分子生物学所要解决的基本科学问题。

作为分子生物学分支的医学分子生物学，所要探讨的基本科学问题，将是上述分子生物学基本问题在医学领域的延伸。她的发展将为医学的发展提供重要的理论和技术基础。具体来讲，医学分子生物学主要解决两方面的问题：一是在理论方面，阐明疾病和亚健康状态发生和发展的分子机制。对这种机制的解释将不会是单一分子、单一结构、单一通路、单一疾病的方式，而是多层次知识的系统整合。二是在解决理论问题的基础上，发展新型分子生物学技术，为疾病的诊断、治疗和预防提供崭新的、可行的、符合人类经济和社会发展需求的手段。

二、医学分子生物学发展过程中的重要事件

回顾医学分子生物学发展过程中的一些重要发现，一方面将有助于我们整理科学发展思路，从中理解医学分子生物学的关键理论和技术发展策略，更好地把握今后的发展趋势；另一方面，更为重要的是充分感谢那些为医学分子生物学的发展做出努力和贡献的科学家，他们的贡献将使人类永远受益。

我们只选择了那些确实属于医学分子生物学范畴的事件，而没有包括其他一些尽管在生命科学中意义更为重大，但并非直接有关医学问题的发现，例如 DNA 双螺旋结构的

阐明。

1956 年第一次将一种疾病的发病机制确认为一个分子的一个特定改变。镰状细胞贫血的细胞学改变在 1910 年由芝加哥医生 E. Iron 报道。1940 年，L. Pauling 曾发现该病的病因是血红蛋白结构改变所致，但是未能确定变化的部位。1956 年，V. Ingram 发现珠蛋白第 6 位氨基酸由谷氨酸突变为缬氨酸是镰状细胞贫血的致病原因。

1959 年鉴定出了第一个染色体病。Down 综合征在 1866 年被首次报道。1959 年，J. Lejeune 和他的同事发现该综合征是由于 21 号染色体三体突变引起的。同年，还有 Turner 综合征和 Klinefelter 被鉴定为性染色体异常。

1961 年第一次对新生儿代谢缺陷作出诊断。医生兼细菌学家 R. Guthrie 建立了检验新生儿是否患有苯丙酮酸尿症的新方法。如果新生儿的苯丙酮酸尿症能够在出生时即被诊断，就可以通过调整饮食结构避免含苯基丙氨酸的氨基酸的摄入，最终可以消除几乎所有的临床症状。

1972 年第一个重组 DNA 分子在实验室制造成功。美国斯坦福大学 P. Berg 带领的研究组在实验室制造出了第一个人工重组 DNA 分子。他们的工作第一次证明了具有无限前景的基因重组技术的可行性。如今临床上正在使用或正在研究的基因工程药物都始于这一重要的技术成果。

1976 年发现癌基因。H. Varmus 和 M. Bishop 从 1970 年开始进行肿瘤病毒学研究。他们从鸡肉瘤中分离到在体外能使鸡胚成纤维细胞转化，在体内能使鸡患肉瘤的病毒。比较研究后，发现了病毒癌基因 *src*。后来，他们又确立了细胞癌基因的概念。

1978 年美籍华裔科学家简悦威（YW. Kan）第一次利用 DNA 多态性与致病基因的关联性，成功地对镰状细胞贫血进行了产前诊断，开创了基因诊断技术在临床应用的新时代。

1981 年第一个转基因小鼠（生长激素）构建成功，开创了建立疾病模型的工作基础。

1982 年世界上第一个基因重组产品——人胰岛素问世。DNA 重组技术的创始人之一 H. Boyer 和风险投资人 R. Swanson 在 1976 年建立了世界上第一家生物技术公司——Genentech 公司。在 1977 年，该公司在细菌中表达了第一个人的蛋白质。1982 年第一个 DNA 重组药物——人胰岛素上市。

1983 年第一次在染色体定位致病基因。亨丁顿病的基因被定位于第 4 号染色体。该病为常染色体显性遗传，引起神经元死亡，中年出现进行性特异性的运动震颤，身体僵硬，痴呆等综合征。染色体定位 10 年后，在 6 个研究小组共 58 名研究者的共同努力下才克隆出致病基因 *huntingtin*，并确切定位在 4p16. 3。

1986 年第一个人类致病基因的定位克隆。慢性肉芽肿病（CGD）为单基因遗传病，以炎症损害为特征。波士顿儿童医院的研究组在不知道基因编码产物的情况下完成了基因的定位（Xp21）和克隆。这是第一个通过定位克隆技术分离得到人的致病基因。同年，杜氏肌营养不良和视网膜母细胞瘤的致病基因也利用定位克隆策略被获得。

1987 年第一张含 400 个 RFLP 标志的人遗传图谱绘制完毕。

1990 年，第一例真正意义上的基因治疗是用腺苷脱氨酶（ADA）基因来治疗 *ADA* 基因缺陷引起的严重复合型免疫缺陷症的患者。正常 *ADA* 基因借助反转录病毒载体被导入患儿的 T 淋巴细胞，再回输给患儿，5 年后患儿体内仍有 10%的造血细胞呈 ADA 基因阳性，这一治疗病例的成功使得基因治疗迅速开展起来。

1995 年完成流感病毒基因组（1830 137 bp）序列测定，这是第一个完成的病原微生物基因组测序，也是第一个完整生物的基因组测序。同年 10 月，支原体基因组 580 070 bp 测序完成。1998 年，结核杆菌基因组测序完成，全长 4 400 000 bp。

2003 年，人类基因组计划序列图完成。

三、医学分子生物学的现状与未来

医学分子生物学目前正处于一个发展十分迅速的时期，最主要的原因是人类基因组研究计划的实施已经彻底改变了以往科学研究的模式。规模化、整体化、自动化、信息化的趋势似乎势不可挡地将涵盖包括医学分子生物学在内的所有生命科学相关领域。事实上，这一趋势的形成是医学面临的人类疾病的复杂性和分子生物学进展到一定程度所带来的必然结果。

随着医学分子生物学研究工作的深入，人们发现对一个或几个基因进行的个别研究几乎不可能解决像肿瘤发病机制、疾病易感性、药物敏感性差异等复杂性问题，因而人类基因组计划应运而生。如今，人类基因组学的序列分析图已经完成，功能基因组学和蛋白质组学已经起步，医学分子生物学的发展方向也将随之产生新的变化。预期未来一段时间，医学分子生物学将可能在以下几个方面取得较多的进展。

首先，对于各种疾病发生发展的分子机制的认识会有相当大的进步。这些进步将体现在对于疾病发生的遗传学背景将有更多和更深入的认识，从基因的突变、基因的多态性和个体基因组与环境相互作用的角度，都会改变目前的很多认识。尤其是在肿瘤、心血管系统疾病、糖尿病、重要感染性疾病等方面的新认识，对于这些疾病为何发生？哪些分子影响疾病的进程？哪些人群更易患病？为什么人们对同样的药物会有不同的敏感性等医学问题，都将从分子及其相互作用的复杂网络层次有新的理解。

其次在诊断方面，基于个体与群体基因组的信息分析将改变目前疾病诊断的现状，在充分认识疾病发生、发展机制的基础上，新的诊断方法不仅将更为准确地对疾病的存在状态给予评价，更将为疾病的转归、预后及疾病在个体发生的风险性做出预测。

在疾病的治疗方面，医学分子生物学的研究手段，如疾病基因组学、药物基因组学、疾病蛋白质组学、疾病动物模型、RNA 干扰技术等将为药物的发现和开发提供更多的新靶点，也将提供更多的药物筛选技术平台。同样，个体基因组学信息将评价和给出个体的最佳用药方案。新基因、新蛋白的结构、功能及相互作用将不断得到阐明，其中的一些将成为基因工程药物和疫苗的候选分子。基因治疗无疑是医学分子生物学另一个诱人领域，基因替代将为许多目前无法医治的疾病提供机会。最后，细胞分化和器官形成的分子机制的阐明可能使未来的器官置换成为常规。

实现上述理想，仍然需要人类几十年甚至几百年的努力。所有从事分子生物学和医学研究及实践的人都将有机会接受挑战，赢得成功。

（药立波）

三、医学分子生物学的现状与未来

第一篇　分子生物学基本原理

基因的活动是分子水平生命活动的核心内容，涉及核酸和蛋白质的结构与功能、基因组的结构与功能、基因的复制与表达、基因表达的调控及其生物学效应，涉及生物大分子之间的相互作用以及这些相互作用所形成的细胞间通讯和细胞内信号转导。认识这些生命活动的分子特征，是理解疾病发生分子机制的基础，应用这些特征，将使疾病的诊断、治疗和预防发生根本变化。

什么是基因？基因的化学结构和生理功能是什么？不同生物种属的基因组在结构上有何不同？基因组如何在细胞分裂时忠实地复制到子代细胞内？基因的损伤和变异是如何发生的？基因如何在细胞内表达并保持特有的时空表达谱？这些分子生物学的最基本的规律性问题将在本篇前6章中予以讨论。同属本教材系列的《生物化学》（五年制本科生用，人民卫生出版社）一书已经从化学角度详细地叙述了核酸的结构与功能以及DNA复制、转录和翻译过程中的化学问题。本书则侧重从基因的结构、基因的组织以及功能的角度讨论上述问题，在一些具体的化学反应过程方面没有放很多的笔墨。另外，本篇所涉及的分子生物学基本原理没有限于人体，而是较系统介绍了原核生物、真核生物和病毒的各自特点，这主要是基于对病毒和致病细菌在医学上的重要性的考虑。

此外，对于基因组学、功能基因组学和蛋白质组学等在医学未来发展中具有重要价值的系统科学问题在本篇做了基本介绍。

环境信息如何通过细胞内的遗传信息控制细胞的基因表达及增殖、分化、发育；细胞接收环境信息的主要方式、关键分子、信号转导通路及其在细胞的增殖分化、基因表达、代谢等生理过程中的调节作用等作为细胞分子生物学的重要部分，是医学分子生物学的重要内容，也将在本篇的最后两章讨论。本部分有关细胞信号转导的内容将侧重介绍信号转导的基本体系和基本的工作方式，并归纳讨论一些规律性的问题，一些具体的信号转导通路，如胰岛素、血管生长因子、趋化因子、发育信号通路等将在后续章节涉及具体问题时再予以介绍。

第一章　基因的结构与功能

关键知识点

- 基因是细胞和个体世代遗传信息贮存和传递的基本单位；基因的化学本质是DNA（RNA病毒除外）；一个基因是DNA分子中具有特定核苷酸排列顺序的一个区段。
- DNA分子的一级结构指碱基的线性排列顺序；双链反向互补的右手螺旋是DNA的主要二级结构；DNA需要进一步折叠为超螺旋结构，在真核细胞内可形成核小体进而折叠为染色体。
- 每个基因都携带两部分遗传信息：可以在细胞内表达为各种RNA（mRNA又可继续指导蛋白质合成）的结构基因序列和为表达这些结构基因所需要的启动子、增强子等调控序列。真核生物的绝大部分结构基因由内含子和外显子间隔排列而成，存在编码序列断裂。
- 不同生物的基因的组成结构各有特点。原核生物的几个结构基因常成串分布在一个操纵子内受一个或数个启动子及其调控序列的共同控制；真核生物基因的每个结构基因都有自己的调控序列。
- 各种RNA是重要的基因表达产物。其中mRNA是蛋白质合成的模板；tRNA作为各种氨基酸的载体并将其转呈给mRNA；rRNA与核糖体蛋白共同构成细胞合成蛋白质的场所。RNA组学研究各种细胞内具有重要功能的非编码小RNA的种类、结构和功能。

关键词：基因；DNA；DNA双螺旋结构模型；核小体；染色体；结构基因；操纵子；内含子；外显子；mRNA；tRNA；rRNA

生物性状的遗传规律早在19世纪60年代已经被奥地利遗传学家G. Mendel所认识。1909年，丹麦生物学家W. Johannsen根据希腊文“给予生命”之义，创造了基因（gene）一词，代替了G. Mendel的“遗传因子”，但是基因的本质和机制在20世纪后半叶的分子生物学时代才真正得以认识。

基因的现代生物学概念是：基因决定遗传性状的表达，基因存在于染色体及线粒体DNA上，呈直线排列并世代相传。

基因的现代分子生物学概念是：基因是核酸分子中贮存遗传信息的遗传单位，是RNA和蛋白质相关遗传信息的基本存在形式，是指贮存RNA序列信息和有功能的蛋白质多肽链序列信息，以及表达这些信息所必需的全部核苷酸序列。大部分生物中构成基因的核酸物质是DNA，少数生物（如RNA病毒）中是RNA。

第一节　基因的化学结构

遗传信息的载体是核酸。核酸（nucleic acid）是以核苷酸为基本组成单位的生物信息大分子，可分为脱氧核糖核酸（deoxyribonucleic acid，DNA）和核糖核酸（ribonucleic acid，RNA）两大类。DNA存在于细胞核和线粒体内，携带遗传信息，决定着细胞和个体的基因型（genotype）；RNA存在于细胞质、细胞核和线粒体内，参与遗传信息的复制与表达。某些RNA病毒则以RNA作为遗传信息的载体。

一、DNA 和 RNA 的化学组成

DNA 是由四种脱氧核糖核苷酸通过 3′,5′-磷酸二酯键连接而成的长链大分子。这 4 种脱氧核糖核苷酸分别是脱氧腺嘌呤核苷酸(deoxyadenosine monophosphate,dAMP)、脱氧鸟嘌呤核苷酸(deoxyguanosine monophosphate,dGMP)、脱氧胞嘧啶核苷酸(deoxycytidine monophosphate,dCMP)、脱氧胸腺嘧啶核苷酸(deoxythymidine monophosphate,dTMP)。由于 4 种核苷酸间的差异主要是碱基不同,因此 DNA 的核苷酸序列可以碱基序列表示。DNA 分子中的碱基是腺嘌呤(adenine,A)、鸟嘌呤(guanine,G)、胞嘧啶(cytosine,C)、胸腺嘧啶(thymine,T),4 种碱基以千变万化的排列组合形成复杂、大量的遗传信息。除 RNA 病毒以外,大多数生物的遗传信息都是以特定的核苷酸排列顺序贮存在 DNA 分子上。

RNA 是由 4 种核糖核苷酸（AMP、GMP、CMP、UMP）通过 3′,5′-磷酸二酯键相连而成的多核苷酸长链。

二、DNA 的一级结构意指分子中的核苷酸排列顺序

DNA 的分子结构可分为一级、二级和三级结构来描述。一级结构（primary structure）是指 DNA 分子中所有脱氧核糖核苷酸的排列顺序，基因对遗传信息的携带和传递是依靠这种核苷酸排列顺序而实现的。

组成 DNA 的脱氧核苷酸中成分相同的脱氧核糖和磷酸基团是构成长链的骨架结构，所以碱基顺序实际上代表了核苷酸的排列顺序。DNA 的一级结构中脱氧核苷酸或核苷酸的连接具有严格的方向性，由前一位脱氧核苷酸的 3′-OH 与下一位脱氧核苷酸的 5′位磷酸基之间形成 3′，5′-磷酸二酯键，从而构成一个没有分支的线性大分子—多聚脱氧核苷酸链（图 1-1）。DNA 的两个末端分别称为 5′末端（游离磷酸基）和 3′末端（游离羟基）。

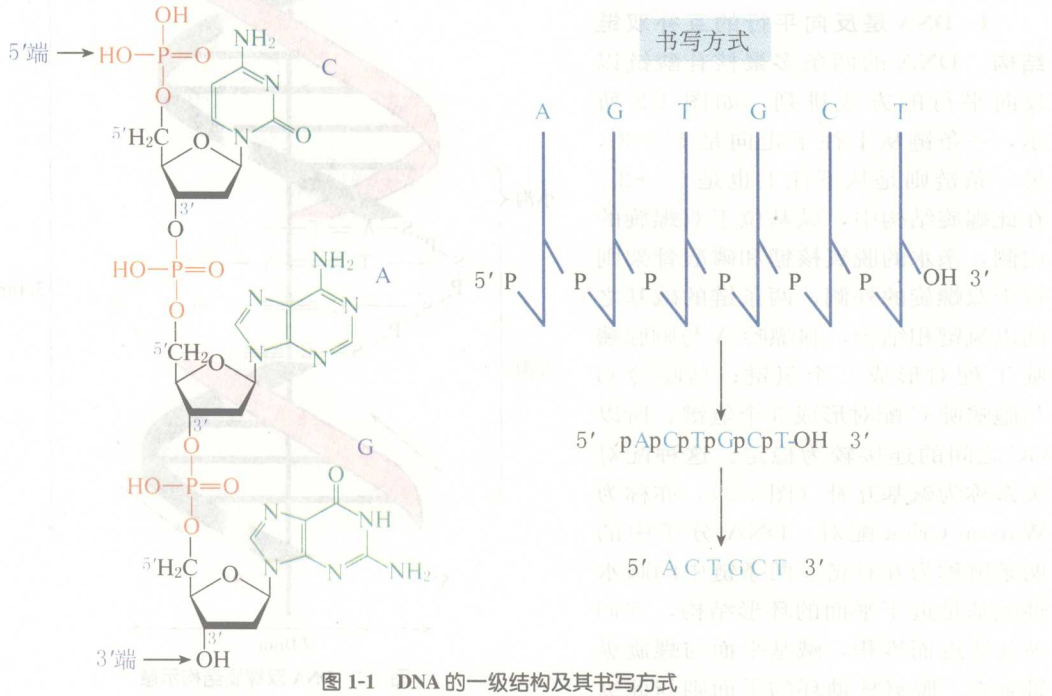


图 1-1　DNA 的一级结构及其书写方式

DNA 一级结构的书写方式可有多种，从简到繁如图 1-1 所示。需要强调的是 DNA 的书写一般应从 5′末端到 3′末端。RNA 的书写方式与 DNA 相同。

三、双链反向互补的右手螺旋是 DNA 的主要二级结构

DNA 分子的二、三级结构指的是分子中所有原子在三维空间具有的相对位置，即 DNA 分子的高级结构。1953 年，J. Watson 和 F. Crick 两人在 Nature 杂志上发表了 DNA 双螺旋（double helix）结构模型的论文。这一发现揭示了生物界遗传性状得以世代相传的分子机制，是分子生物学发展的里程碑。迄今 55 年，这一历史性发现已经为生物学和医学的发展带来了巨大的变化，无论是对生命活动本质的深刻理解，还是改造生命的各种努力，无一不得益于这一发现所揭示的自然奥秘。

（一）Chargaff 规则是最重要的 DNA 双螺旋结构的研究基础

20 世纪中期，人们已经证实了 DNA 是遗传信息的载体，阐明 DNA 的分子结构很快成为当时最为引人注目的科学问题之一。E. Chargaff 等人对此做出了重要贡献。他们采用层析和紫外吸收分析等技术分析了 DNA 分子的碱基成分，提出了关于 DNA 分子中 4 种碱基组成的 Chargaff 规则：①腺嘌呤与胸腺嘧啶的摩尔数相等，而鸟嘌呤与胞嘧啶的摩尔数相等；②不同生物种属的 DNA 碱基组成不同；③同一个体不同器官、不同组织的 DNA 具有相同的碱基组成。这一规则提示了 DNA 分子中存在着碱基配对，分别是 A 与 T，G 与 C 配对。1953 年，Watson 和 Crick 综合当时人们对于 DNA 分子特性的各种认识和获得的各种数据，并根据 R. Franklin 拍摄的高质量 DNA 分子 X 线衍射照片，提出了 DNA 分子双螺旋（double helix）的结构模型，故也称为 Watson-Crick 模型。

（二）DNA 双螺旋结构有三个要点

1. DNA 是反向平行的互补双链结构　DNA 的两条多聚核苷酸链以反向平行的方式排列，如图 1-2 所示，一条链从上往下走向是 5′→3′，另一条链则是从下往上也是 5′→3′。在此螺旋结构中，碱基位于双螺旋的内侧，亲水的脱氧核糖和磷酸骨架则位于双螺旋的外侧，两条链的碱基之间以氢键相结合，腺嘌呤 A 与胸腺嘧啶 T 配对形成 2 个氢键；鸟嘌呤 G 与胞嘧啶 C 配对形成 3 个氢键。所以 GC 之间的连接较为稳定。这种配对关系称为碱基互补（图 1-3），亦称为 Watson-Crick 配对。DNA 分子中的两条链称为互补链。两条链中的疏水性碱基是近于平面的环形结构，它们彼此接近而堆积，碱基平面与螺旋纵轴垂直，脱氧核糖环的平面则与此纵

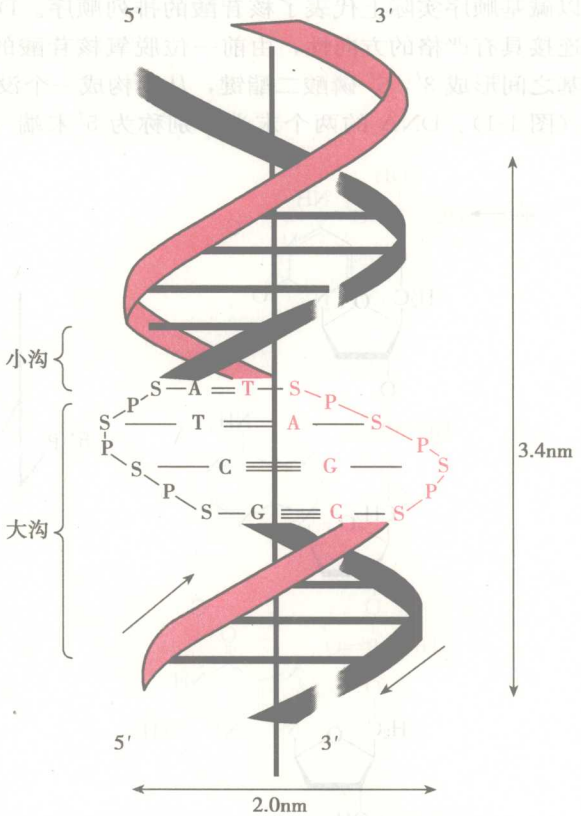


图 1-2　DNA 双螺旋结构示意

轴平行。对于双链 DNA 长度可用碱基对（base pair，bp）、千碱基对（kilo-base pair，kbp）作单位来描述。

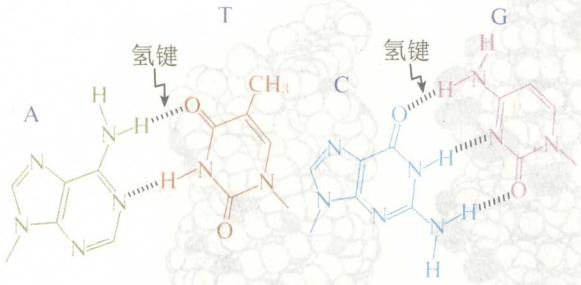


图 1-3　DNA 双螺旋结构中的碱基配对

2. DNA 分子是右手双螺旋　DNA 的两条脱氧多聚核苷酸链反向平行围绕同一中心轴旋转，呈右手螺旋。两个相邻的碱基对之间距离为 0.34 nm，两个碱基之间有 36°的旋转角度。沿中心轴每旋转一周有 10 个核苷酸，螺距为 3.4 nm，螺旋的直径是 2.0 nm。外观上，DNA 双螺旋分子表面存在一个大沟（major groove）和一个小沟（minor groove）。这些沟状结构与相关蛋白质因子对其的识别有关。

3. 疏水性碱基堆积力和氢键是 DNA 双螺旋结构的稳定力　DNA 双链结构的横向稳定依靠两条链互补碱基间的氢键维系，纵向则靠碱基平面的堆积力维持，碱基堆积力是指螺旋内部碱基平面重叠靠近产生的疏水作用。从总能量意义上来讲，碱基堆积力对于双螺旋的稳定性更为重要。

（三）DNA 双螺旋结构具有多样性

Watson-Crick 结构模型是从生理盐溶液中抽出 DNA 纤维，在 92%相对湿度下作 X 射线衍射图谱分析而获得的，这是 DNA 分子在水性环境和生理条件下最稳定的结构。后来人们发现 DNA 的双螺旋结构在溶液的离子强度或相对湿度改变时，其表面沟的深浅、螺距、旋转角度等都会发生一些变化。1979 年，A. Rich 等人在研究人工合成的 CGCGCG 的晶体结构时发现这种 DNA 是左手螺旋。后来又发现天然 DNA 分子中同样存在左手螺旋。目前人们将 Watson-Crick 模型的 DNA 结构称为 B-DNA。将 Rich 等人的 DNA 双螺旋结构称为 Z-DNA。另外还有 A-DNA 的存在（图 1-4）。

DNA 双螺旋中的核苷酸除了前述 A/T，G/C 之间氢键外，还能形成一些附加氢键，如：另一个 T 与 A/T 碱基对的 A 之间，可形成额外的 2 个氢键，使得这 3 个碱基形成了 T * A/T 配对；当 pH 降低时，胞嘧啶的 N-3 可以质子化，质子化的胞嘧啶与 G/C 碱基对的 G 又可形成 2 个氢键，3 个碱基形成了 C^+ * G/C 的配对。这种配对是 K. Hoogsteen 在 1963 年发现的，因此称为 Hoogsteen 配对。由于碱基之间有形成 Hoogsteen 配对的可能，因此可以形成三螺旋 DNA(triplex DNAs)；在某些人工合成的寡核苷酸链中也观察到了这种形式的 DNA。

可见，B-DNA 的结构不是自然界 DNA 的唯一存在方式。在生物体内，不同构象的 DNA 在功能上可能有所差异，与基因表达的调节和控制相适应。

四、DNA 在细胞内组装为致密结构

DNA 是一个长度很大的生物大分子，如人体细胞内的全部 DNA 的总长度约为 3×10^9 bp，如果连在一起可达 1 米多。显然，DNA 的长度要求其必须进一步折叠成为更致密的结构才能够存在于很小的细胞核内。实际上，DNA 在细胞内以双螺旋为结构基础进一步旋转折叠形成超螺旋（supercoils）结构。盘绕方向与 DNA 双螺旋方同相同为正超螺旋（positive supercoil）；盘绕方向与 DNA 双螺旋方向相反则为负超螺旋（negative supercoil）。自然界的闭合双链 DNA 主要是以负超螺旋形式存在。

1. 绝大部分原核生物的 DNA 都是共价封闭的环状双螺旋分子　原核生物的 DNA 在

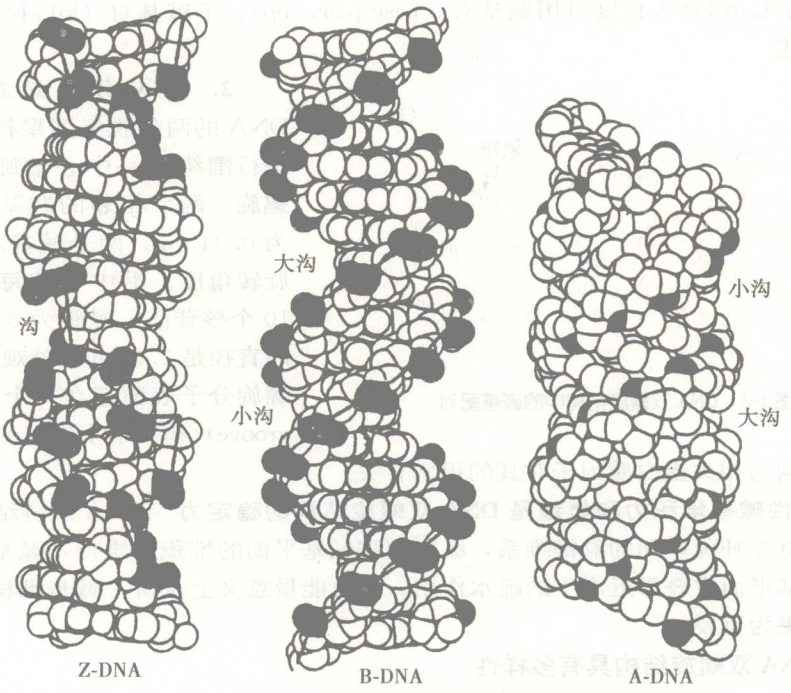


图 1-4　DNA 双螺旋结构的多样性

细胞内进一步盘绕，并形成类核（nucleoid）结构，以保证其以较致密的形式存在于细胞内。类核结构中约 80%为 DNA，其余为蛋白质。在细菌基因组中，超螺旋可以相互独立存在，形成超螺旋区（图 1-5），各区域间的 DNA 可以有不同程度的超螺旋结构。目前的分析表明，大肠杆菌的基因组中，平均每 200 个碱基就有一个负超螺旋形成。

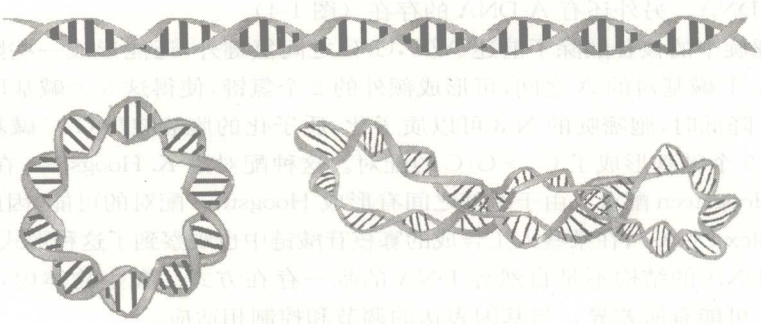


图 1-5　线性、环状和超螺旋 DNA 示意图

2. 真核生物 DNA 在细胞核内可组装成染色体　真核生物 DNA 与蛋白质形成的复合物以非常致密的形式存在于细胞核内，基本结构单位是核小体（nucleosome）。在细胞周期的大部分时间里以分散存在的染色质（chromatin）形式出现，在细胞分裂期形成高度组织有序的染色体（chromosome），在光学显微镜下即可见到。

核小体由 DNA 和 5 种组蛋白（histone）共同构成。组蛋白是小分子量的碱性蛋白质，富含精氨酸和赖氨酸。核小体中的组蛋白共有 5 种：H1、H2A、H2B、H3 和 H4。各两分子的 H2A、H2B、H3 和 H4 构成八聚体的核心组蛋白，双链 DNA 缠绕在这一核心上形成核小体的核心颗粒（core particle）。核小体的核心颗粒之间再由 DNA（约 60 bp）和组蛋白 H1 构成的连接区相连形成串珠样的结构（图 1-6）。

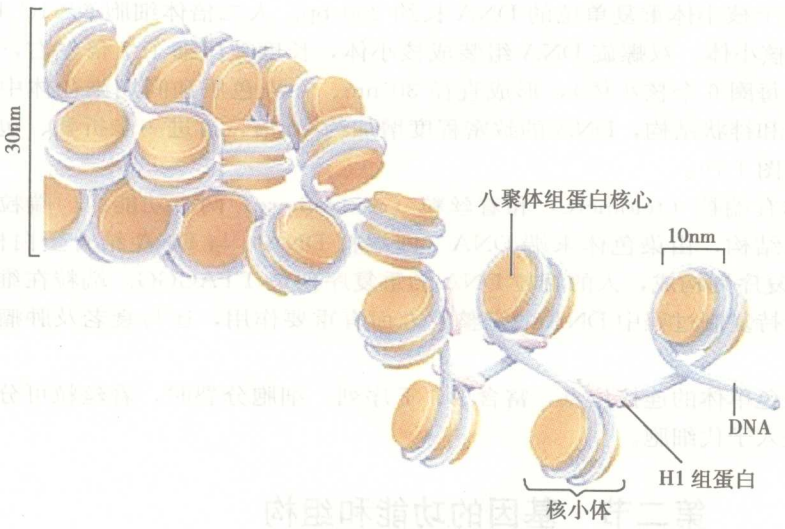


图 1-6　核小体结构示意图

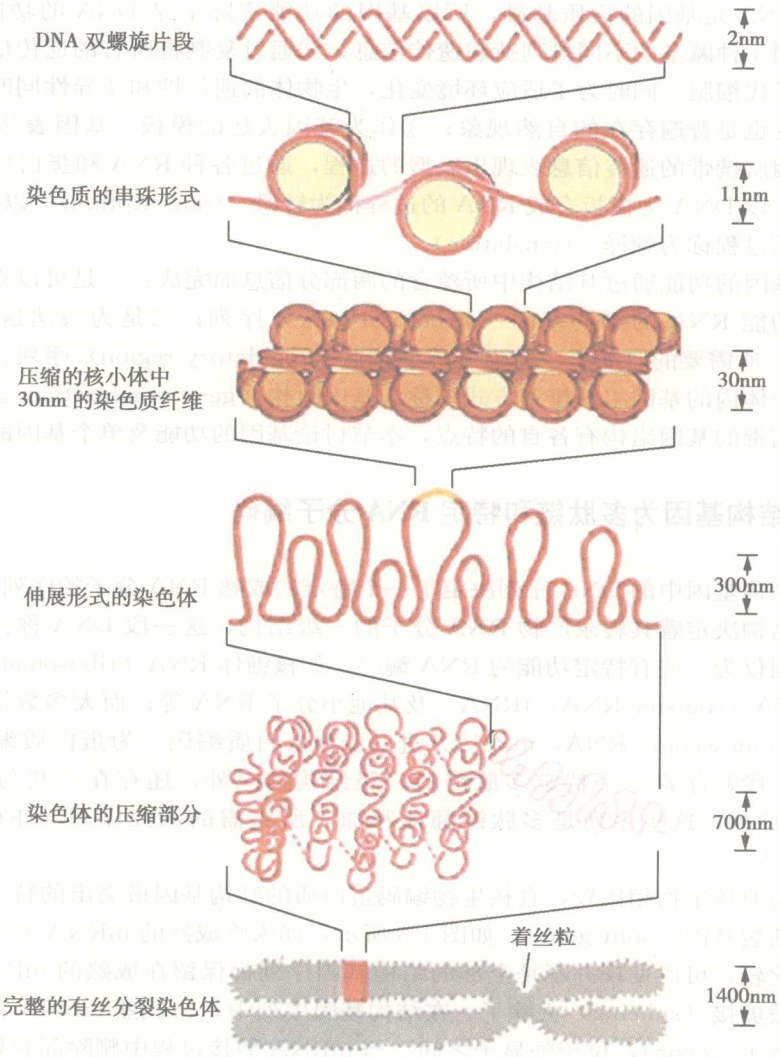


图 1-7　核小体折叠形成染色体

串珠样结构中每个核小体重复单位的 DNA 长约 200 bp。人二倍体细胞 6×10^9 bp 的 DNA 含约 3×10^7 个核小体。双螺旋 DNA 组装成核小体，长度被压缩了 7 倍左右，核小体进一步折叠卷曲（每圈 6 个核小体），形成直径 30 nm、在染色质和间期染色体中都可以见到的纤维状结构和袢状结构，DNA 的致密程度增加约 40 倍，再进一步折叠、包装即为染色质和染色体（图 1-7）。

真核生物染色体有端粒（telomere）和着丝粒（centromere）两个功能区。端粒是染色体末端膨大的粒状结构，由染色体末端 DNA（即端粒 DNA）与 DNA 结合蛋白构成。端粒 DNA 由简单重复序列构成，人的端粒 DNA 的重复序列是 TTAGGG。端粒在维持染色体结构的稳定和维持复制过程中 DNA 的完整性方面有重要作用，还与衰老及肿瘤的发生发展有关。

着丝粒是两个染色单体的连接位点，富含 A、T 序列。细胞分裂时，着丝粒可分开使染色体均等有序地进入子代细胞。

第二节　基因的功能和组构

DNA 是基因的物质基础，所以基因的功能实际上是 DNA 的功能。基因的功能是：①利用 4 种碱基的不同排列荷载遗传信息；②通过复制将所有的遗传信息稳定、忠实地遗传给子代细胞。同时为了适应环境变化，生物体的遗传性和变异性同时存在，变异性即基因突变也是普遍存在的自然现象；③作为基因表达的模板。基因表达（gene expression）是基因所携带的遗传信息表现出表型的过程，通过各种 RNA 和蛋白质在细胞内有序合成实现。以 DNA 为模板合成 RNA 的过程称为转录（transcription）；以 RNA 为模板合成蛋白质的过程称为翻译（translation）。

基因的功能通过其结构中所蕴含的两部分信息而完成：一是可以在细胞内表达为蛋白质或功能 RNA 的结构基因（structural gene）序列；二是为表达这些结构基因（合成 RNA）所需要的启动子、增强子等调控区（regulatory region）序列。单个基因的组成结构及个体内的基因组织排列方式统称为基因组构（gene organization）。原核生物、真核生物和病毒的基因组构有各自的特点，本节讨论基因的功能及单个基因的组成结构。

一、结构基因为多肽链和特定 RNA 分子编码

结构基因中的 DNA 序列决定了一个特定的成熟 RNA 分子的序列，换句话说 DNA 的一级结构决定着其转录产物 RNA 分子的一级结构，这一段 DNA 称为结构基因。有的结构基因仅为一些有特定功能的 RNA 编码，如核糖体 RNA（ribosomal RNA，rRNA）、转运 RNA（transfer RNA，tRNA）及其他小分子 RNA 等；而大多数结构基因则通过信使 RNA（messenger RNA，mRNA）进一步为蛋白质编码。为蛋白质编码的结构基因中除含有一段贮存着一个特定多肽链的一级结构信息外，还存在一些与多肽链编码无关的 DNA 序列，这些序列是多肽链翻译和加工所必需的调控序列，亦称为 mRNA 的非翻译区。

与原核生物相比较，真核生物编码蛋白质的结构基因最突出的特点是其不连续性，被称为断裂基因（split gene）。如图 1-8 所示，如果将成熟的 mRNA 分子序列与其结构基因序列比较，可以发现并不是全部的结构基因序列都保留在成熟的 mRNA 分子中，有一些区段被剪接（splicing）去除了。在结构基因序列中，与成熟 mRNA 分子相对应的序列称为外显子（exon）；位于外显子之间、与 mRNA 剪接过程中删除部分相对应的序列则称为

内含子（intron）。外显子与内含子相间排列，共同组成结构基因。每个基因的内含子的数目比外显子要少 1 个。内含子和外显子同时出现在最初合成 mRNA 前体中，在合成后被剪接（见第五章）。如全长为 7.7 kbp 的鸡卵清蛋白基因有 8 个外显子和 7 个内含子，最初合成的 RNA 前体与相应的结构基因是等长的，内含子序列被切除后的成熟 mRNA 分子的长度仅为 1.2 kbp。在不同的结构基因中外显子的数量不同，少则数个，多则数十个。外显子的数量是描述基因结构的重要特征之一。

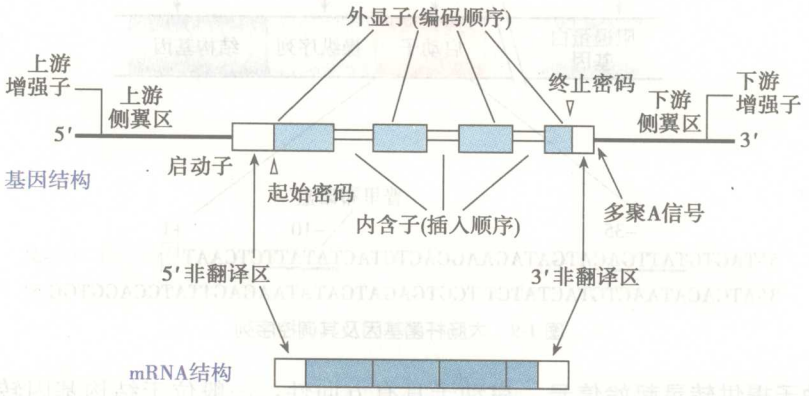


图 1-8　真核生物的结构基因

高等真核生物内绝大部分编码蛋白质的结构基因都有内含子，只有组蛋白编码基因是一例外。此外，编码 rRNA 和个别 tRNA 的结构基因也都有内含子。低等真核生物的内含子分布差别很大，有的酵母的结构基因较少有内含子，有的则较常见。原核生物的结构基因基本没有内含子。病毒的结构基因常与宿主基因的结构特征相似。感染细菌的病毒（噬菌体）的基因与细菌基因的结构特征相似，结构基因是连续的；而感染真核细胞的病毒的基因具有某些真核生物基因结构特征，少量的结构基因也由于含有内含子而间断。病毒结构基因由于基因组大小的限制，有的还存在着重叠编码，更为有效地利用基因序列。

二、基因包含结构基因和调控序列

单个基因的组成结构中除了结构基因的编码序列以外，还包括对结构基因的表达起调控作用的区域，如启动子、增强子等调节区。对一个基因的完整描述不仅针对它的编码区，同时也包括它的调控区。一个基因的调控区和结构基因位于同一个 DNA 分子中的相邻部位，这种调节方式称为顺式调节（*cis*-regulation），相应的 DNA 序列称为顺式调控元件或称顺式作用元件（*cis*-acting element）。顺式作用元件的作用均需要通过结合相应的酶或蛋白因子方可实现，而这些蛋白质一般由位于另外的 DNA 分子或同一 DNA 分子的远距离部位的基因来编码，因而被称为反式作用因子（*trans*-acting factors）。反式作用因子是通过直接结合或间接作用于 DNA、对基因表达发挥不同作用（激活或抑制）的各类蛋白因子。

为方便叙述调节序列和编码序列的关系，人们约定将一个基因的 5′端称之为上游，3′端称为下游；为标定 DNA 信息的具体位置，将结构基因上开始 RNA 链合成的第一个核苷酸所对应的碱基记为＋1，在此碱基上游的序列记为负数，向 5′端依次为－1、－2、…；在此碱基下游的序列记为正数，向 3′端依次为＋2、＋3、…。零不用于标记碱基位置。

（一）原核生物基因的调控序列

原核生物基因的转录调控序列中，最基本的是启动子（promoter）和转录终止信号，在不同的基因中尚有可被其他调节蛋白（阻遏蛋白或激活蛋白）所识别和结合的顺式作用元件。图 1-9 是一原核生物基因的典型结构组成及其序列特征。

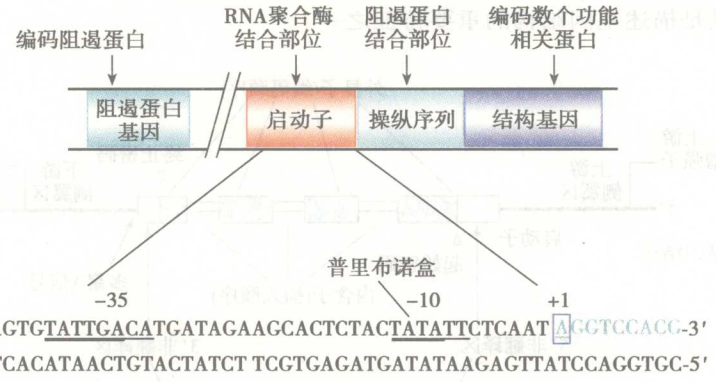


图 1-9　大肠杆菌基因及其调控序列

1. 启动子提供转录起始信号　启动子具有方向性，一般位于结构基因转录起始点的上游。不同基因间的启动子序列上存在一定保守性，有一致性序列（consensus sequences）或称共有序列存在。启动子序列本身不出现于 RNA 产物中，仅提供转录起始信号。大肠杆菌的启动子区的长度为 40～60 bp，至少包括了三个功能区：一是 RNA 合成的起点，即＋1 位碱基；二是位于－10 bp 区的 RNA 聚合酶结合部位，有着“TATAAT”的共有特征序列，亦称为普里布诺盒（Pribnow box）；三是转录起始识别部位（recognition site），位于－35 bp 区，共有序列是“TGACA”。尽管存在着上述共有序列，但原核生物的启动子间序列可有较大差异。启动子的序列越接近共有序列，起始转录的作用越强，称为强启动子；反之为弱启动子。例如 λ 噬菌体的 *PL*、*PR* 以及 T7 噬菌体的 P_{T7} 是强启动子，而乳酸操纵元的 P_{Lac} 是较弱的启动子。

2. 终止子提供 RNA 合成终止信号　在结构基因下游近 3′端的一段 DNA 序列中有 GC 富集区组成的具有回文特征的重复序列（见第五章），转录后在 RNA 分子中形成特殊的结构以终止 RNA 链的延伸。

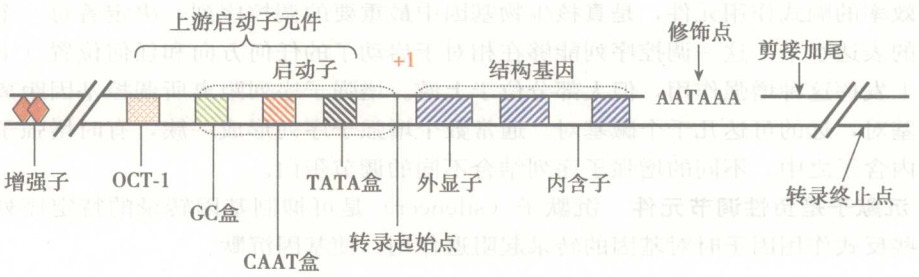
3. 操纵元件被阻遏蛋白识别与结合　操纵元件或称操纵序列（operator）是一些启动子邻近部位的一小段特定序列，可被具有抑制转录作用的阻遏蛋白识别并结合，通常与启动子区域有部分重叠。

4. 正调控蛋白结合位点可加强下游结构基因的转录　在前已述及的原核基因的弱启动子附近常有一些特殊的 DNA 序列，某些具有转录激活作用的正调控蛋白可以识别并结合这种 DNA 序列，加快转录的启动。

（二）真核生物基因的调控序列

真核生物基因的调控序列远较原核生物复杂，迄今了解仍很有限。真核生物顺式作用元件包括启动子、上游调控元件、增强子、加尾信号和一些细胞信号反应元件等（图 1-10）。

1. 启动子提供转录起始信号　大部分真核细胞基因的启动子位于结构基因转录起始点的上游，启动子本身通常不被转录；但有一些启动子（如编码 tRNA 基因的启动子）的 DNA 序列可以位于转录起始点的下游，这些 DNA 序列可以被转录。真核生物主要有三类



OCT-1: ATTTGCAT八聚体

图 1-10　真核生物基因的一般结构

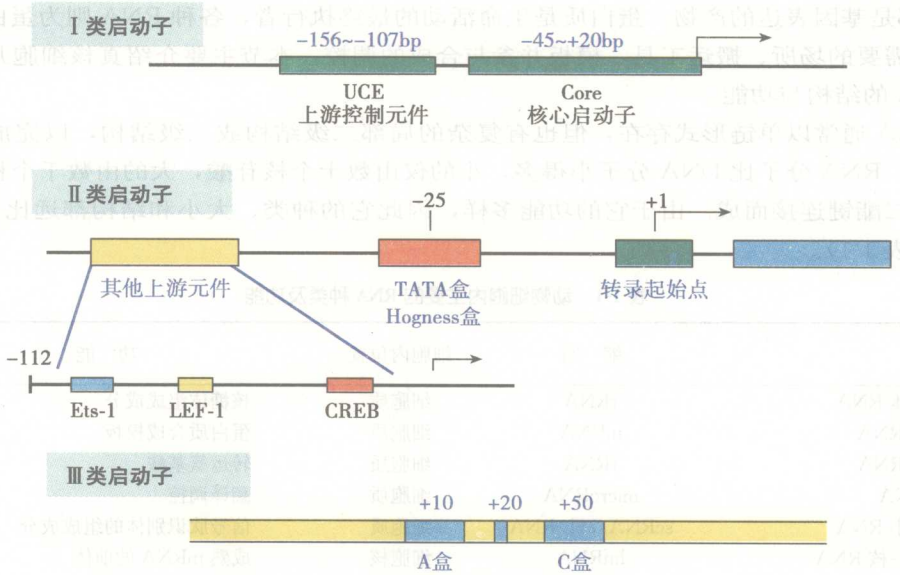


图 1-11　真核生物基因的三类启动子

启动子（图 1-11），分别对应于细胞内存在的三种不同的 RNA 合成酶和相关蛋白质。

（1）Ⅰ类启动子富含 GC 碱基对：具有Ⅰ类启动子的基因主要是编码 rRNA 的基因。Ⅰ类启动子包括核心元件（core element）和上游调控元件（upstream control element，UCE）两部分，前者位于－45 bp 到 ＋20 bp，转录起始的效率很低，后者位于－156 bp～－107 bp，能增强转录的起始。两部分序列都富含 GC 碱基对。

（2）Ⅱ类启动子具有 TATA 盒特征结构：具有Ⅱ类启动子的基因主要是编码蛋白质（mRNA）的基因和一些小 RNA 基因。Ⅱ类启动子通常是由 TATA 盒（TATA box）、几个上游启动子元件和转录起始位点组成。TATA 盒位于转录起始点上游－25 bp 处，其核心序列是 TATA（A/T）A（A/T），决定着 RNA 合成的起始位点。有的Ⅱ类启动子在 TATA 盒的上游还可存在 CAAT 盒、GC 盒等特征序列。共同组成启动子。

（3）Ⅲ类启动子包括 A 盒、B 盒和 C 盒：具有Ⅲ类启动子的基因包括 5S rRNA、tRNA、U6 snRNA 等 RNA 分子的编码基因。Ⅲ类启动子的位置较独特，如 tRNA 基因的启动子有 A 盒、B 盒和 C 盒三部分，分别位于＋10 bp～＋20 bp 和＋50 bp～＋60 bp 两个区域。

2. 增强子增强邻近基因的转录　增强子（enhancer）是可以增强真核生物基因启动子

的工作效率的顺式作用元件，是真核生物基因中最重要的调控序列，决定着每一个基因在细胞内的表达水平。这一调控序列能够在相对于启动子的任何方向和任何位置（上游或者下游）上发挥这种增强作用，但大部分位于上游。增强子序列距离所调控基因距离近者几十个碱基对，远的可达几千个碱基对。通常数个增强子序列形成一簇，有时增强子序列也可位于内含子之中。不同的增强子序列结合不同的调节蛋白。

3. 沉默子是负性调节元件　沉默子（silencer）是可抑制基因转录的特定序列，当其结合一些反式作用因子时对基因的转录起阻遏作用，使基因沉默。

第三节　RNA 的结构与功能

基因的功能需要通过表型来体现，基因表达产物是产生表型的分子基础。蛋白质和 RNA 都是基因表达的产物。蛋白质是生命活动的最终执行者，各种 RNA 则为蛋白质合成提供所需要的场所、搬运工具、模板并参与合成的调控。本节主要介绍真核细胞几种重要的 RNA 的结构与功能。

RNA 通常以单链形式存在，但也有复杂的局部二级结构或三级结构，以完成一些特殊功能。RNA 分子比 DNA 分子小得多，小的仅由数十个核苷酸，大的由数千个核苷酸通过磷酸二酯键连接而成。由于它的功能多样，因此它的种类、大小和结构都远比 DNA 多样化（表 1-1）。

表 1-1　动物细胞内主要的 RNA 种类及功能

	缩　写	细胞内位置	功　能
核糖体 RNA	rRNA	细胞质	核糖体组成成分
信使 RNA	mRNA	细胞质	蛋白质合成模板
转运 RNA	tRNA	细胞质	转运氨基酸
微 RNA	microRNA	细胞质	翻译调控
胞质小 RNA	scRNA/7SL-RNA	细胞质	信号肽识别体的组成成分
不均一核 RNA	hnRNA	细胞核	成熟 mRNA 的前体
核内小 RNA	snRNA	细胞核	参与 hnRNA 的剪接、转运
核仁小 RNA	snoRNA	核仁	rRNA 的加工和修饰
线粒体核糖体 RNA	mt rRNA	线粒体	核糖体组成成分
线粒体信使 RNA	mt mRNA	线粒体	蛋白质合成模板
线粒体转运 RNA	mt tRNA	线粒体	转运氨基酸

一、信使 RNA 具有特征结构

20 世纪 50 年代中期，DNA 序列决定蛋白质序列已经得到了公认，但是如何解释真核生物的 DNA 存在于细胞核，而蛋白质合成是在细胞质中进行这一事实？直到 1960 年 F. Jacob 和 J. Monod 等人发现了 mRNA 是蛋白质合成的模板才使基因组的遗传信息如何为蛋白质编码这一生命科学的基本问题得以解决。

1. 真核生物和原核生物的 mRNA 中都有编码区和非翻译区　mRNA 的功能是转录基因组 DNA 遗传信息的碱基排列顺序，并指导蛋白质合成中的氨基酸排列顺序。在 mRNA 分子中，中间的一部分序列是一个特定多肽链的序列信息，这一段核苷酸序列称为多肽链编码区或开放阅读框（open reading frame，ORF），此段核苷酸序列决定着多肽链分子的一级结构。ORF 通常从 mRNA 分子 5′端的第一个 AUG 开始，每 3 个核苷酸为一组，决

定肽链上一个氨基酸，称为三联体密码（triplet code）或密码子（codon），直至终止密码子结束。在开放阅读框的5′端上游和3′端下游的核苷酸序列没有编码功能，称为非翻译区（untranslated region，UTR）。非翻译区的功能主要是参与翻译调控。

2. 真核生物mRNA特有的5′末端的帽结构　大部分真核细胞的mRNA的5′末端以7-甲基鸟嘌呤-三磷酸鸟苷为起始结构（图1-12），这种m^7GpppN结构被称为帽结构（cap sequence）。帽结构中的鸟苷酸及相邻的A或G都可以发生甲基化，由于甲基化位置的差别可产生数种不同的帽结构。帽结构可以与一类称为帽结合蛋白（cap binding proteins，CBPs）的分子结合。这种mRNA和CBPs复合物对于mRNA从细胞核向细胞质的转运、与核糖体的结合、与翻译起始因子的结合、mRNA稳定性的维系等均有至关重要的意义。

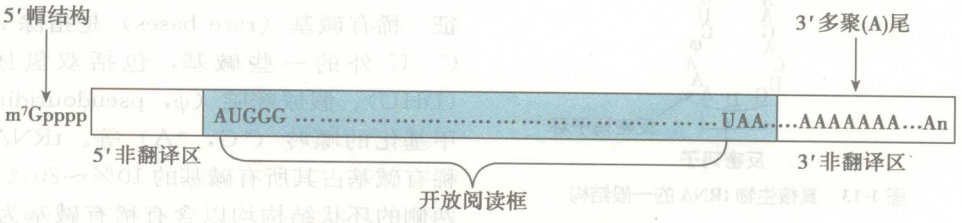


图1-12　真核生物mRNA的一般结构

3. 真核生物mRNA的3′末端有多聚A尾结构　在真核生物mRNA的3′末端，大多数有一段由数十个至百余个腺苷酸连接而成的多聚腺苷酸结构，称为多聚A尾［poly（A）］。poly（A）结构也是在mRNA转录完成以后额外加入的。Poly（A）在细胞内与poly（A）结合蛋白［poly（A）-binding protein，PABP］相结合而存在，每10～20个碱基结合一个PABP单体。所以，真核细胞的mRNA的3′端实际上是一个poly（A）和蛋白多聚体形成的复合物。目前认为，这种3′末端多聚A尾结构和5′帽结构共同负责mRNA从核内向胞质的转位、mRNA的稳定性维系以及翻译起始的调控。去除多聚A尾和帽结构是细胞内mRNA降解的重要步骤。

生物体各种mRNA链的长短差别很大，主要是由其转录的模板DNA区段的大小及转录后的剪接方式所决定的。mRNA分子的长短决定了它要翻译出的蛋白质的分子量大小。在各种RNA分子中，mRNA的半衰期最短，由几分钟到数小时不等。

二、转运RNA有氨基酸接纳茎和反密码环

各种转运RNA均由74～95个核苷酸组成，序列中存在着一些能局部互补配对的区域，可以形成局部的双链，中间不能配对的部分则膨出，因而形成特有的茎环（stem-loop）结构（图1-13），使得tRNA整个分子的形状类似于三叶草形（cloverleaf pattern）。茎环结构，例如氨基酸接纳茎和反密码环就是tRNA在蛋白质合成过程中作为各种氨基酸的载体、并将其转呈给mRNA的结构基础。

1. 氨基酸接纳茎位于tRNA分子的3′末端　所有tRNA的3′端的最后3个碱基均为CCA，是氨基酸的结合部位，称为氨基酸接纳茎（acceptor stem）。不同tRNA的氨基酸接纳茎结合不同的氨基酸。有的氨基酸只有一种tRNA作为载体，有的则需要数种tRNA作载体，这是由于有的氨基酸可有多种密码子为其编码，即存在着密码子的简并性（详见基因表达一章）。不同tRNA的命名采用右上标的不同氨基酸的三字母代号，如$tRNA^{Tyr}$表示可以与酪氨酸相结合的tRNA。已经结合了氨基酸的tRNA用前缀氨基酸三字母代号表示，如Tyr-$tRNA^{Tyr}$代表$tRNA^{Tyr}$的氨基酸接纳茎已经结合有酪氨酸。

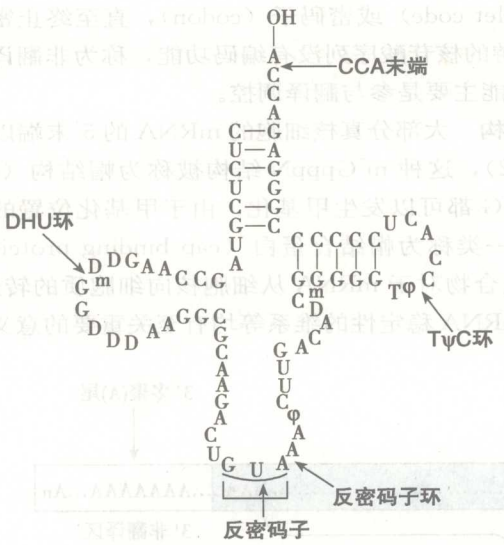


图 1-13　真核生物 tRNA 的一般结构

2. 反密码子位于 tRNA 分子反密码环　每个 tRNA 分子中都有连续排列的 3 个碱基与 mRNA 上编码相应氨基酸的密码子具有碱基互补关系，可以配对结合，这 3 个碱基被称为反密码子（anticodon），位于反密码环内。不同的 tRNA 有不同的反密码子，这是辨认 mRNA 上相互补的密码子、将其所携带的氨基酸正确安放在延伸肽链上的结构基础。

3. 含有稀有碱基是 tRNA 分子的结构特征　稀有碱基（rare bases）是指除 A、G、C、U 外的一些碱基，包括双氢尿嘧啶（DHU）、假尿嘧啶（ψ，pseudouridine）和甲基化的嘌呤（^{m}G，^{m}A）等。tRNA 中的稀有碱基占其所有碱基的 10%～20%。位于两侧的环状结构均以含有稀有碱基为特征，分别称为 DHU 环和 TψC 环。

X 射线衍射结构分析表明，tRNA 的共同三级结构是倒 L 型。tRNA 的 TψC 环与 DHU 环在三叶草形的二级结构上各处一方，但在三级结构上相距很近。对于这种空间结构的具体功能目前还缺乏详细的认识，不过这些结构是 tRNA 与核糖体中的蛋白质和 rRNA 的相互作用密切相关。

三、核糖体 RNA 是核糖体的组成成分

核糖体 RNA 是细胞内含量最多的 RNA，约占 RNA 总量的 80%以上。rRNA 与核糖体蛋白（ribosomal protein）共同构成核糖体（ribosome）或称为核蛋白体。

核糖体是 rRNA 和蛋白质形成的超级大分子复合物（表 1-2），原核生物和真核生物的核糖体均由易于解聚的大、小两个亚基组成。原核生物核糖体的沉降系数为 70S（S 是大分子物质在超速离心沉降中的一个物理学单位，可间接反映分子量的大小），由 30S 小亚基与 50S 大亚基组成；真核生物核糖体的沉降系数为 80S，由 40S 小亚基与 60S 大亚基组成。核糖体位于细胞质中。

表 1-2　核糖体的组成

	原核生物（以大肠杆菌为例）		真核生物（以小鼠肝为例）	
小亚基		30S		40S
rRNA	16S	1542 个核苷酸	18S	1874 个核苷酸
蛋白质	21 种	占总重量的 40%	33 种	占总重量的 50%
大亚基		50S		60S
rRNA	23S	2940 个核苷酸	28S	4718 个核苷酸
	5S	120 个核苷酸	5.85S	160 个核苷酸
			5S	120 个核苷酸
蛋白质	31 种	占总重量的 30%	49 种	占总重量的 35%

四、小分子 RNA 有不同的种类和功能

细胞核、核仁及细胞质中含有许多小分子 RNA，它们的长度一般小于 300 个核苷酸，也是通过转录后加工而来，具有各种重要的生理功能。这些小 RNA 被通称为非 mRNA 小 RNA（small non-messenger RNA，snmRNA）。对于 snmRNA 的研究近年来得到重视，由此产生的 RNA 组学（RNomics）就是专门研究细胞中 snmRNA 的种类、结构和功能的前沿学科。

1. 核小 RNA 参与 mRNA 的剪接　真核细胞 RNA 前体中的内含子剪接需要核小 RNA（small nuclear RNA，snRNA）。由于 snRNA 序列中尿嘧啶含量较高，因此又用 U 命名。其中 U1、U2、U4、U5 和 U6 位于核浆内，参与 mRNA 的剪接；U3 存在于核仁，与 rRNA 加工有关；snRNA 需要与相应蛋白形成核小核糖核蛋白颗粒（small nuclear ribonucleoprotein particle，snRNP）方可发挥作用。

2. 核仁小 RNA 参与 rRNA 前体的剪接　核仁小 RNA（small nucleolar RNA，snoRNA）广泛存在从单细胞生物酵母至哺乳动物等。snoRNA 也可与特定的蛋白质结合形成 snoRNP，在核仁区富集，参与 rRNA 加工以及核糖体的装配。

3. 胞质小 RNA 参与蛋白质的翻译和转运　在细胞质中也有一些具有特定功能的小分子 RNA，被称为胞质小 RNA（small cytoplasmic RNA，scRNA），如 7SL RNA 等。scRNA 参加蛋白质翻译合成后的转运。

4. 小 RNA 参与基因表达调控　微 RNA（microRNA，miRNA）、小干扰 RNA（small interfering RNA，siRNA）、反义 RNA（antisense RNA）是一些在基因表达调控，尤其是转录后调控中发挥重要作用的小 RNA。它们的作用将在基因表达调控一章中介绍。

第四节　核　酸　酶

核酸酶是指所有可以水解核酸的酶。依据底物是 DNA 或是 RNA，分别称为 DNA 酶（deoxyribonuclease，DNase）和 RNA 酶（ribonuclease，RNase）。有的核酸酶可以在 DNA 或 RNA 分子内部切断磷酸二酯键，称为核酸内切酶（endonuclease）。在核酸内切酶中，有的只水解双链分子，有的只水解单链分子；有的要求序列特异性（4～8 bp），称为限制性核酸内切酶（restriction endonuclease）；有的没有序列特异性。核酸外切酶（exonuclease）仅能水解位于核酸分子链末端的核苷酸。根据其作用的方向性，又有 5′→3′或 3′→5′核酸外切酶之分。

核酶（ribozyme）是一类具有催化作用的 RNA，其底物亦是核酸，因此从功能上来讲也属于核酸内切酶，且为序列特异性的核酸内切酶。除了其在 RNA 合成后的加工修饰作用机制正在逐步阐明外，核酶也正在作为针对病毒或肿瘤基因的药物而被广泛研究，有的已经进入临床试验阶段。除了催化性 RNA 可以作为序列特异性的核酸内切酶降解 mRNA 外，人工合成的寡聚脱氧核苷酸片段也具有序列特异性降解 RNA 的作用，称为催化性 DNA（DNAzyme）。催化性 DNA 与催化性 RNA 相比，具有更好的化学稳定性和生物学稳定性，在疾病治疗方面的将有更好的前景。目前尚未发现天然的催化性 DNA 的存在。

（张　英）

第二章　基因组的结构与功能

关键知识点

- 细胞或生物体中，一套完整单倍体的遗传物质的总和称为基因组。基因组的结构主要指不同的基因功能区域在核酸分子中的分布和排列情况，基因组的功能是贮存和表达遗传信息。不同的生物体，其基因组的大小和复杂程度各不相同，进化程度越高的生物其基因组越复杂。
- 真核生物的基因组庞大，除了结构基因的结构与原核生物大不相同以外，还存在大量的非结构基因区域。真核细胞基因组具有结构基因呈断裂基因形式、以单顺反子转录、结构基因在基因组中所占比例远小于非编码区域的特点。真核细胞基因组存在着大量重复序列，可以分为高度重复序列、中度重复序列和单拷贝或低度重复序列等 3 种。真核基因组的另一结构特点是存在多基因家族。人的染色体基因组 DNA 长约 3.0×10^9 bp，约编码 28 000 个基因。人的线粒体 DNA 全长 16 569 bp，编码 37 个基因。
- 原核生物染色体基因组中结构基因的编码特点是结构基因连续编码、编码序列不重叠、很少有重复序列。原核生物的结构基因与调控序列以操纵子的形式存在。在操纵子结构中，数个功能上有关联的结构基因串联在一起，共同构成信息编码区。信息编码区的上游是调控区。质粒是细菌细胞内的、染色体外的 DNA 分子，能够独立于细胞的染色质 DNA 而进行复制。
- 病毒的基因组较小。大多数 DNA 病毒的基因组是双链 DNA 分子，而大多数 RNA 病毒的基因组是单链 RNA 分子。病毒基因组有基因重叠编码现象。

关键词： 基因组；重复序列；多基因家族；操纵子；质粒

细胞或生物体的一套完整单倍体遗传物质的总和称为基因组（genome），1920 年由德国科学家 H. Winkles 首先使用，用“genes”和“chromosomes”两个词组合来描述生物的全部基因和染色体。人类基因组包含了细胞核染色体 DNA（常染色体和性染色体）及线粒体 DNA 所携带的所有遗传物质。不同的生物体基因组的大小和复杂程度各不相同。

表 2-1 列出了从原核生物到真核生物有代表性的生物体中 DNA 分子的大小。病毒、原核生物以及真核生物所贮存的遗传信息量有着巨大的差别，其基因组的结构与组织形式上也各有特点。本章重点讨论的是与医学密切相关的人类基因组和某些细菌与病毒基因组的结构，包括：①基因组层面的基因组构，主要指基因组中的基因的组织排列方式；②基

表 2-1　不同生物体中 DNA 的比较

	大小（kbp）	双螺旋线性长度（m）	染色体数（单倍体）	形　状
原核生物				
病毒 SV40	5.2	0.000 001 7	1	环状
噬菌体 φX174	5.4	0.000 001 8	1	线状单链
噬菌体 λ	46	0.000 015	1	线状
细菌大肠杆菌	4 000	0.001 36	1	环状
真核生物				
酵母	1 000	0.000 34	17	线状
果蝇	41 000	0.014	4	线状
人	3 000 000	1.20	24	线状

因组中的基因种类、数目和分布。基因组的总体功能是贮存和表达遗传信息，其具体机制已经分别在基因、基因组复制、基因表达及其调控等各章中具体介绍，这里仅涉及特征性的基因组结构的功能意义。

第一节　人基因组的结构与功能

全球生命科学的大协作项目-人类基因组计划的完成将我们对人类基因组结构与功能的认识带到了新的层次。人类基因组包含了位于细胞核内的染色体 DNA 和位于线粒体内的线粒体 DNA，其组成成分如图 2-1 所示。人类基因组中只有很少一部分的 DNA 序列用以编码蛋白质。

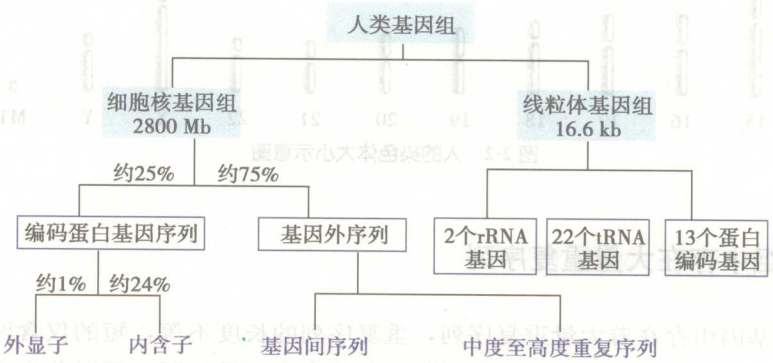


图 2-1　人类基因组的组成

一、人基因组中只有不到三万种基因

人的遗传信息由 DNA 携带，细胞核内的单倍体基因组由 3.0×10^9 bp 组成，按平均 1 000 个碱基编码 1 种蛋白质计，理论上可有 300 万个基因。在发现基因组 DNA 中大量存在非编码序列后，仍然推测人的基因数目在 10 万左右。人类基因组计划的初步完成使基因的推测数目减少到 3.4 万～3.5 万个，目前基于基因组序列测定的基因数目约 28 900 个（截至 2008 年 2 月）。这些基因在人的常染色体及性染色体上（图 2-2）的分布见表 2-2。

表 2-2　人的核染色体上的基因数量分布

染色体	碱基数（Mb）	基因数	染色体	碱基数（Mb）	基因数	染色体	碱基数（Mb）	基因数
1	247	2 782	9	140	1 148	17	79	1 469
2	243	1 888	10	135	1 106	18	76	432
3	200	1 469	11	134	1 848	19	64	1 695
4	191	1 154	12	132	1 370	20	62	737
5	181	1 268	13	114	551	21	47	352
6	171	1 505	14	106	1 275	22	50	742
7	159	1 452	15	100	945	X	155	1 336
8	146	984	16	89	1 109	Y	58	307

Mb：兆碱基对；合计：28 924

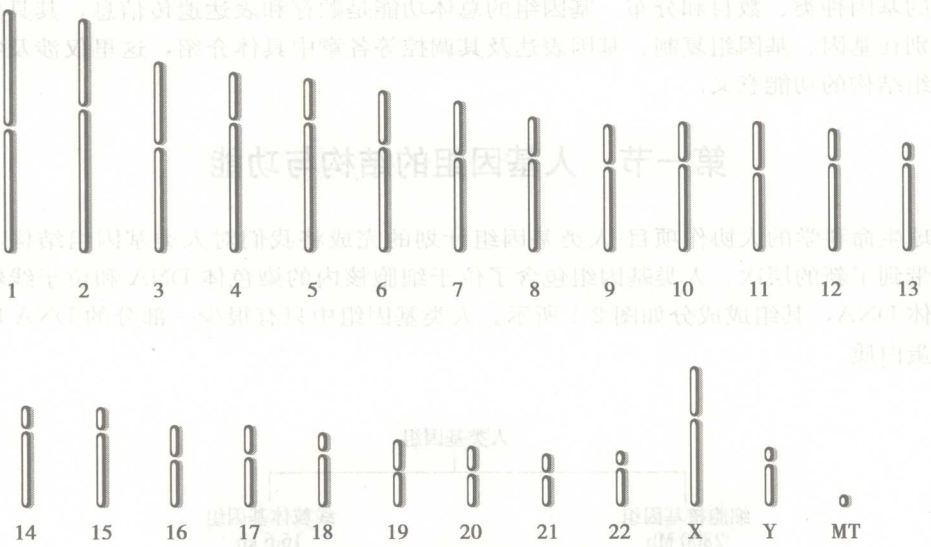


图 2-2　人的染色体大小示意图

二、人基因组中存在大量重复序列

真核细胞基因组存在着大量重复序列，重复序列的长度不等，短的仅含两个碱基，长的多达数百、乃至上千个碱基。重复序列的重复频率也不尽相同，可以分为高度重复序列（highly repetitive sequences）、中度重复序列（moderately repetitive sequences）和单拷贝或低度重复序列（single copy sequences）等 3 种。

（一）高度重复序列

高度重复序列重复频率可达 10^6 次，包括卫星 DNA、反向重复序列和较复杂的重复单位组成的重复序列。在基因组中所占比例随种属而异，从 10%～60%不等。在人基因组中，高度重复序列约占基因组长度的 20%。高度重复序列又按其结构特点分为以下 2 种：

1. 反向重复序列　反向重复序列由两个相同顺序的互补拷贝在同一 DNA 链上反向排列而成，其总长度约占人基因组的 5%。若以两个互补拷贝组成的反向重复为一个单位，则反向重复的单位长度约为 300 bp 或略短。两个单位之间有一平均 1.6 kbp 的片段相隔，两对反向重复单位之间的平均距离约 12 kbp，意味着它们多数是散在，而非群集于基因组中。

2. 卫星 DNA　卫星 DNA（satellite DNA）是另一类高度重复序列，这类重复序列的重复单位一般由 2～10 bp 组成，成串排列。由于这类序列的碱基组成不同于基因组的其他部分，可用等密度梯度离心法将其与主体 DNA 分开，因而称为卫星 DNA 或随体 DNA。在人基因组中卫星 DNA 约占 5%～6%。

目前认为，高度重复序列的功能是：①参与复制水平的调节。反向重复序列常存在于 DNA 复制起点区的附近，是一些蛋白质（包括酶）的结合位点；②参与基因表达的调控。高度重复序列可以转录到核内不均一 RNA 分子中，而有些反向重复序列可以形成发夹结构，有助于稳定 RNA 分子；③参与染色体配对。如 α 卫星 DNA 成簇样分布在染色体着丝粒附近，可能与染色体减数分裂时染色体配对有关，即同源染色体之间的联会可能依赖于具有染色体专一性的特定卫星 DNA 序列。

（二）中度重复序列

中度重复序列指在真核基因组中重复数十至数万（$<10^5$）次的重复序列。少数在基因组中成串排列在一个区域，大多数与单拷贝基因间隔排列。依据重复序列的长度，中度重复序列可分为两种类型：

1. 短分散重复片段　短分散重复片段（short interspersed repeated segments，SINES）这类重复序列的平均长度约为 300 bp～500 bp，与平均长度约为 1 000 bp 的单拷贝序列间隔排列。拷贝数可达 10 万左右。如 Alu 家族，Hinf 家族等属于这种类型的中度重复序列。

Alu 家族是哺乳动物，包括人基因组中含量最丰富的一种短分散片段中度重复序列，平均每 5 kbp DNA 就有一个 Alu 序列，在单倍体人基因组中重复达 30 万～50 万次，约占人基因组的 3%～6%。Alu 家族每个成员的长度约 300 bp，由于每个单位长度中有一个限制性内切酶 *Alu* 的切点（AG↓CT），从而将其切成长 130 bp 和 170 bp 的两段，因而命名为 Alu 序列（或 Alu 家族）。Alu 序列具有种属特异性，根据人的 Alu 序列制备的探针只能用于检测人基因组中的 Alu 序列。人的 Alu 序列是由两个约 130 bp 的正向重复序列构成的二聚体，在第二个单体中有一个 31bp 的插入序列，该插入序列在 Alu 家族的不同成员之间核苷酸序列相似但不相同。每个 Alu 序列两侧为 6～20 bp 的正向重复序列，不同的 Alu 成员的侧翼重复序列也各不相同。Alu 序列的 5′ 端比较保守，但富含脱氧腺苷酸残基的 3′端在不同的 Alu 成员中是有变化的。Alu 序列广泛散布于整个基因组的原因可能是由于 Alu 序列可由 RNA 聚合酶转录成 RNA 分子，再经反转录酶（reverse transcriptase）的作用形成互补 DNA（complementary DNA，cDNA），然后重新插入基因组。

KpnI 家族是中度重复序列中仅次于 Alu 家族的第二大家族。用限制性内切酶 *Kpn* I 消化人类及其他灵长类动物的 DNA，在电泳谱上可以看到 4 个不同长度的片段，分别为 1.2、1.5、1.8 和 1.9 kbp，这就是所谓的 KpnI 家族。KpnI 家族成员呈散在分布，属于中度重复序列的长分散片段。尽管不同长度类型的 KpnI 家族之间同源性比较小，不能互相杂交，但它们的 3′端有广泛的同源性。KpnI 家族的拷贝数约为 3 000～4 800 个，占人类基因组的 1%。与散在分布的 Alu 家族相似，KpnI 家族中有一部份是通过 KpnI 序列的 RNA 转录产物的 cDNA 拷贝重新插入到人基因组 DNA 中产生的。

Hinf 家族以 319 bp 长度的串联重复存在于人基因组中。用限制性内切酶 *Hinf* Ⅰ消化人的基因组 DNA，可以分离到这一片段。Hinf 家族在人单倍体基因组内约有 50 100 个拷贝，分散在不同的区域。

2. 长分散重复片段　长分散重复片段（long interspersed repeated segments，LINES）重复序列的平均长度为 3 500 bp～5 000 bp，与平均长度为 13 000 bp（个别长几万 bp）的单拷贝序列间隔排列。

中度重复序列在基因组中所占比例在不同种属之间差异很大，一般约占 10%～40%，在人约为 12%。这些序列大多不编码蛋白质。这些非编码的中度重复序列的功能可能类似于高度重复序列。在结构基因之间以及内含子内都可以见到中度重复序列。

真核生物基因组中的 rRNA 基因也属于中度重复序列。与其他中度重复序列不同，各重复单位中的 rRNA 基因都是相同的。rRNA 基因通常集中成簇存在，而不是分散于基因组中，这样的区域称为 rDNA 区，如染色体的核仁组织区（nucleolus organizer region）即为 rDNA 区。人类的 rRNA 基因位于 13、14、15、21 和 22 号染色体的核仁组织区，每个核仁组织区平均含有 50 个 rRNA 基因的重复单位；5S rRNA 基因似乎全部位于 1 号染色体（1q42～43）上，每个单倍体基因组约有 1 000 个 5S rRNA 基因。

（三）单拷贝序列（低度重复序列）

单拷贝序列在单倍体基因组中只出现一次或数次，大多数为蛋白质编码的基因属于这一

类。在基因组中，单拷贝序列的两侧往往为散在分布的重复序列。由于某些单拷贝序列编码蛋白质，体现了生物的各种功能，因此对这些序列的研究对医学实践有特别重要的意义。

三、多基因家族与假基因

真核基因组的另一结构特点是存在多基因家族（multigene family）。多基因家族是指由某一祖先基因经过重复和变异所产生的一组基因。多基因家族大致可分为两类：一类是基因家族成簇地分布在某一条染色体上，它们可同时发挥作用，合成某些蛋白质，如组蛋白基因家族就成簇地集中在第 7 号染色体长臂 3 区 2 带到 3 区 6 带区域内；另一类是一个基因家族的不同成员成簇地分布于不同染色体上，这些不同成员编码一组功能上紧密相关的蛋白质，如珠蛋白基因家族。一个多基因家族中可有多个基因，根据结构与功能的不同又可以分为亚家族（subfamily），例如人的低分子量 G 蛋白家族至少有 50 多个成员，其中又进一步分为 *RAS*、*RAC* 等亚家族。

人的基因组中存在假基因（pseudogene）。假基因是基因组中存在的一段与正常基因非常相似但不能表达的 DNA 序列。这类基因可能曾经有过功能，但在进化中获得的一个或几个突变，造成了序列上的细微改变从而阻碍了正常的转录和翻译功能，使它们不再能编码 RNA 和蛋白质产物。与相应的正常基因相比，假基因往往缺少正常基因的内含子，两侧有顺向重复序列。人们推测，假基因的来源之一，可能是基因经过转录后生成的 RNA 前体通过剪接失去内含子形成 mRNA，如果 mRNA 经反复反转录产生 cDNA，再整合到染色体 DNA 中去，便有可能成为假基因，因此该假基因是没有内含子的，在这个过程中，可能同时会发生缺失，倒位或点突变等变化，从而使假基因不能表达。

四、线粒体 DNA 结构有别于染色体 DNA

线粒体是细胞内的一种重要细胞器，是生物氧化的场所，一个细胞可拥有数百至上千个线粒体。线粒体 DNA（mitochondrial DNA，mtDNA）可以独立编码线粒体中的一些蛋白质，因此 mtDNA 是核外遗传物质。mtDNA 的结构与原核生物的 DNA 类似，是环状分

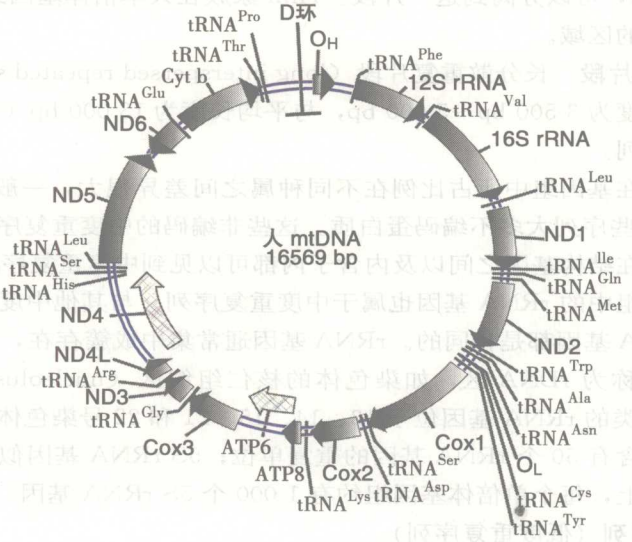


图 2-3　人的线粒体基因组

子。线粒体基因的结构特点也与原核生物相似。

人的线粒体基因组全长 16 569 bp，共编码 37 个基因（图 2-3），包括 13 个编码构成呼吸链多酶体系的一些多肽的基因、22 个编码 mt-tRNA 的基因、2 个编码 mt-rRNA（16S 和 12S）的基因。

第二节　原核生物基因组的结构与功能

原核生物的基因组较小，因此比较容易研究，对其结构和功能的认识也远较真核生物、尤其是高等动植物深入。原核生物基因组 DNA 虽与蛋白结合，但并不形成真正的染色体结构，只是习惯上将之称为染色体 DNA。原核生物的基因组主要是染色体 DNA。某些原核生物还有质粒等其他携带遗传物质的 DNA。

原核生物基因组的结构和功能与真核生物基因组的主要不同点除了结构基因的编码是连续的以外，还存在着以下特点：①基因组中很少有重复序列；②编码蛋白质的结构基因多为单拷贝基因，但编码 rRNA 的基因仍然是多拷贝基因；③结构基因在基因组中所占的比例（约占 50%）远远大于真核基因组，但小于病毒基因组；④许多结构基因在基因组中以操纵子（operon）为单位排列。

一、原核生物基因以操纵子方式组构

原核生物的结构基因与调控序列以操纵子的形式组织在一起。在操纵子结构中，数个功能上有关联的结构基因串联排列，共同构成信息编码区。这些结构基因共用一个启动子和 1 个转录终止信号序列，因此转录合成时仅产生一条 mRNA 长链，为几种不同的蛋白质编码。这样的 mRNA 分子携带了几个多肽链的编码信息，被称为多顺反子 mRNA（polycistron）。一个多顺反子 mRNA 分子转录了一个操纵子内的所有结构基因的信息，所以，原核生物中结构基因的转录调节往往是以操纵子为单元进行的（见第六章）。

一个操纵子至少由启动子、调节这些结构基因表达的操纵元件、几个串联的结构基因区和转录终止信号组成。操纵元件并非结构基因，而是一段能被特异的阻遏蛋白识别和结合的 DNA 序列。启动子是基因转录调控中 RNA 聚合酶和各种调控蛋白作用的部位，是决定基因表达效率的关键元件。

原核生物所完成的生命活动不仅是简单的基因组复制，还要有复杂的代谢活动，即获取外界环境的营养成分，产生能量并合成自身生长所需的材料（核苷酸、氨基酸等）。原核生物需要根据外界环境的变化，调控自身的酶系统的组成及功能，调整细胞内某些蛋白质的数量，适应环境的变化（如高温、渗透压改变等）。因此，原核生物基因组的结构基因数量和功能的类型远远多于病毒基因组。

二、原核生物中的质粒 DNA

质粒（plasmid）是细菌细胞内一种自我复制的环状双链 DNA 分子，能稳定地独立存在于染色体外，并传递到子代，不整合到宿主染色体 DNA 上。质粒的分子量一般为 $10^6 \sim 10^8$，小型质粒的长度一般为 1.5～15 kbp。质粒常含抗生素抗性基因，经过人工改造后的质粒 DNA 是重组 DNA 技术中常用的载体（见第十九章）。

质粒只有在宿主细胞内才能够完成自己的复制，一旦离开宿主就无法复制和扩增。质粒对宿主细胞的生存不是必需的，宿主细胞离开了质粒依旧能够存活。尽管质粒不是细菌

生长、繁殖所必需的物质，但它所携带的遗传信息能赋予细菌特定的遗传性状，如耐药性质粒带有耐药基因，可以使宿主细菌获得耐受相应抗生素的能力；一些人类致病菌的毒力基因亦存在于质粒中，如炭疽杆菌中编码炭疽毒素的基因。

质粒在细菌中的遗传方式有以下特点：①质粒能够在不断增殖的细胞中以一定的拷贝数（copy number）稳定遗传，这是由于质粒自身带有一个复制调控系统，控制质粒复制的频率；②质粒中存在特殊的机制与细菌分裂相适应，使质粒在细菌分裂过程中精确分配（partition）到子细胞中；③质粒中的细胞分裂控制系统能够抑制细胞分裂，使细胞分裂与质粒复制协调，避免太多不含质粒的子代细菌出现，该系统对低拷贝数质粒的稳定遗传十分重要；④质粒在宿主菌中具有不相容性（incompatibility）。具有相同复制起始位点的两种质粒不能共存于一个宿主菌。不过当一个宿主菌中的两个质粒的复制起始位点不同时，它们有各自的分配系统，能精确调节它们在子代细胞中的分配，因此两种质粒可在子代细胞中稳定共存。

第三节　病毒基因组的结构与功能

不同的病毒基因组在大小和结构上有比较大的差异。如乙肝病毒 DNA 只有 3 kbp 大小，只能编码 4 种蛋白质；而痘病毒的基因组有 300 kbp，可以编码几百种蛋白质，不但可以为病毒复制所涉及的酶类编码，而且为核苷酸代谢的酶类编码，因此，痘病毒对宿主的依赖性较乙肝病毒小得多。病毒基因组的特征可以归纳为以下几点：

（一）病毒的基因组较小

与真核细胞和细菌基因组相比，病毒基因组要小得多。即使是痘病毒的基因组有 300 kb，仍然不到大肠杆菌基因组大小的十分之一。

（二）有的病毒基因组是 DNA，而有的病毒基因组是 RNA

每种病毒中只含有一种核酸，或为 DNA 或为 RNA，两者不共存于同一病毒中。基因组为 DNA 的病毒称为 DNA 病毒，基因组为 RNA 的病毒则称为 RNA 病毒。组成病毒基因组的 DNA 或 RNA 可以是单链结构，也可以是双链结构；可以是闭环分子，也可以是线性分子。如乳头瘤病毒是一种闭环的双链 DNA 病毒，而腺病毒的基因组则是线性的双链 DNA；脊髓灰质炎病毒是一种单链的 RNA 病毒，而呼肠孤病毒的基因组是双链的 RNA 分子。大多数 DNA 病毒的基因组是双链 DNA 分子，而大多数 RNA 病毒的基因组是单链 RNA 分子。

（三）RNA 病毒基因组可以由数条不相连的 RNA 链组成

病毒基因组 RNA 可以由不相连的几条 RNA 链构成，如流感病毒的基因组 RNA 分子由 8 个 RNA 分子构成，每个 RNA 分子都含有编码蛋白质分子的信息；呼肠孤病毒的基因组由 10 个不相连的双链 RNA 片段构成，同样每段 RNA 分子都编码一种蛋白质。截至目前，尚未在 DNA 病毒中发现类似现象。

（四）病毒基因组存在着基因重叠

病毒基因组大小十分有限，因此在进化过程中形成了基因重叠编码现象，即同一 DNA 片段可以是 2 种甚至 3 种蛋白质分子的部分编码区，换句话说，两种不同蛋白质的编码区有一部分是共用的。这样的重叠编码机制提高了病毒基因组的编码能力。重叠基因是 1977 年研究 ΦX174 噬菌体时发现的。ΦX174 感染大肠杆菌后共计可合成 11 个蛋白质分子，总分子量为 25 万左右，相当于 6 078 个核苷酸所容纳的信息量。而该病毒 DNA 本身只有 5 375 个核苷酸，最多能编码总分子量为 20 万的蛋白质分子，由此发现了基因重叠现象。

病毒基因组不仅存在着结构基因的重叠，也存在着编码区和调控区的重叠。乳头瘤病毒是一类感染人和动物的病毒，基因组约 8.0 kbp，其中不翻译的部分约为 1.0 kbp，该区同时作为其他基因表达的调控区。

（五）有的病毒基因可以形成多顺反子 mRNA

病毒基因组中有的功能相关的基因或 rRNA 基因，可以形成一个转录单元。它们可被一起转录成为多顺反子 mRNA，然后再加工成为各种蛋白质编码的模板 mRNA。如腺病毒晚期基因编码病毒的 12 种外壳蛋白，转录时是在一个启动子的作用下生成的多顺反子 mRNA，然后再加工成各种 mRNA，编码病毒的各种外壳蛋白。

（六）噬菌体基因是连续的，而真核细胞病毒的基因是不连续的

除了正链 RNA 病毒之外，感染真核细胞的病毒的基因都是先转录成 mRNA 前体，再经加工才能切除内含子成为成熟的 mRNA。由于基因重叠现象，某一基因的内含子或其中的一部分，可以是另一个基因的外显子。

尽管目前人们对于基因组的结构与功能已经有了相当多的认识，人类基因组计划（见本书第七章）的初步目标也已经实现，但是要全面阐明基因组的结构，尤其是完全破解人类基因组全部序列中所蕴藏着的信息和生物学意义，仍然需要艰难而长久的努力。

（德　伟）

第三章　基因组复制

关键知识点

- 基因组 DNA 的复制是 DNA 聚合酶催化的、以亲代 DNA 为模板、dNTP 为底物的化学反应。拓扑异构酶、解旋酶、单链 DNA 结合蛋白、引发酶和 DNA 连接酶等多种分子为该反应所必需。DNA 聚合酶对模板和配对碱基的识别作用及其对错配碱基的校正作用保证了复制的精确性。
- 基因组 DNA 以半保留方式复制，从固定的起点双向进行，需要先合成 RNA 引物，双链中的一条链（前导链）连续合成、另一条链（随后链）分段合成，这种半不连续的复制方式可以克服 DNA 空间结构对 DNA 新链合成的制约。
- 大肠杆菌有 3 种 DNA 聚合酶，DNA 聚合酶Ⅲ是 DNA 复制的主要酶。复制从在复制起点 *ori*C 处解链、形成引发体并合成 RNA 引物开始。DNA 聚合酶Ⅲ二聚体同时催化领头链和随后链的合成。
- 真核生物的 DNA 复制时需要松解核小体结构，复制后又重新组装。复制在多个复制点上同时开始，需要数种 DNA 聚合酶及其他多种蛋白质分子共同完成。DNA 聚合酶 α 在 RNA 引物上启动 DNA 合成，然后由具有高速延伸能力的 DNA 聚合酶 δ 取代继续进行链的延伸。复制中染色体 DNA 末端的端粒结构需要端粒酶催化形成。
- 反转录病毒 RNA 的复制由反转录酶催化。反转录酶的作用是以 RNA 为模板指导三磷酸脱氧核苷酸合成 cDNA。
- 真核生物和原核生物的 DNA 复制都需要严格与细胞周期进程协调进行。复制起始相关因子的浓度和复制起点的甲基化修饰是原核细胞的协调机制；细胞周期蛋白和复制许可因子可能参与真核细胞的细胞周期中复制的控制。

关键词： 复制叉；双向复制；半不连续复制；半保留复制；引发体；冈崎片段；复制子；端粒

细胞正常功能的维持依赖于基因组遗传信息的完整，因此基因组信息在细胞分裂时必须精确地得到复制。基因组信息的准确复制及其与细胞分裂的协调有赖于细胞内一套完善而复杂的 DNA 复制和复制调控体系的存在（RNA 病毒遗传信息有其独特方式复制）。所有生物的基因组 DNA 的复制都遵循共同的规律，不同生物物种因基因组结构上的差异又有各自的复制特点。本章拟先对基因组 DNA 复制的共同规律做简要的归纳和总结，建立起 DNA 复制的整体概念，然后再分别讨论原核生物、真核生物、病毒等基因组的 DNA 复制特点。

第一节　基因组 DNA 复制的基本过程和共同特点

基因组 DNA 复制的过程是以构成基因组的全套 DNA 分子为模板，精确合成一套新的 DNA 分子的细胞内化学反应。多年的研究已经对 DNA 复制的基本过程和规律有所了解，但是对其中很多细节问题尚不能做出明确解释。

一、DNA 复制是酶促化学反应

参与 DNA 复制合成的酶和蛋白因子及其所催化的化学反应特点按其在 DNA 复制过程中出现的先后归纳在表 3-1 中。

表 3-1　参与 DNA 复制合成的主要的酶和蛋白因子

种　类	作　用	作用特点	备　注
解旋酶（helicase）	解开 DNA 双螺旋链	由 ATP 提供能量	在原核生物称为 DnaB 蛋白
单链 DNA 结合蛋白（single-strand DNA binding protein）	特异结合于单链 DNA，稳定复制所要求的单链结构	复制过程中被循环利用	
拓扑异构酶（topoisomerase）	改变 DNA 的拓扑性质，解除其对复制的空间阻碍	兼有内切酶和连接酶两种活性	
引发酶（primase）	属 RNA 聚合酶，负责合成 DNA 复制所需的 RNA 引物	与其他蛋白共同构成复制的引发体	又称引物酶，在原核生物称为 DnaG 蛋白
DNA 聚合酶（DNA polymerase）	负责脱氧多核苷酸的聚合	需以双链模板起始，合成方向为 5′→3′	原核生物和真核生物都有多种，在复制中起不同的作用
DNA 连接酶（DNA ligase）	催化两个 DNA 片段以 3′，5′-磷酸二酯键连接，负责冈崎片段的连接	需要 ATP 提供能量	

二、DNA 复制的基本过程

DNA 的复制如同转录和翻译等过程一样，被人们习惯地分为三个阶段，即起始阶段、延伸阶段和终止阶段。这种划分只是为了叙述上的方便，整个复制过程实际上是连贯不可分的。

1. DNA 复制的起始　复制总是在 DNA 分子上的一个或多个位点开始。这些位点称为复制起点（origin of replication），常简写为 *ori*C。多种蛋白质分子参与 DNA 复制，完成诸如识别复制起始位点、解开 DNA 双链结构、稳定暂时形成的 DNA 局部单链、合成 RNA 引物等功能。在此过程中，这些蛋白将形成 DNA 复制的引发体（primosome）。

2. DNA 链的延伸　DNA 链的延伸主要由 DNA 聚合酶完成。DNA 聚合酶催化的反应是以亲代 DNA 为模板、4 种脱氧核糖核苷三磷酸（deoxynucleotide triphosphate，dNTP），即 dATP、dGTP、dCTP 和 dTTP 为原料底物，依据碱基配对原则，在 RNA 引物链的 3′-OH 上开始逐个添加 dNTP，从而延伸为子代 DNA 链。原核生物有 3 种 DNA 聚合酶，分别称为 DNA 聚合酶Ⅰ、Ⅱ、Ⅲ，真核生物亦有 α、β、γ 等数种 DNA 聚合酶。

复制开始后，解旋酶不断地打开 DNA 双链，复制叉向前推进，导致复制叉前方的 DNA 双螺旋产生反向旋转，形成正超螺旋结构，阻碍复制的继续进行。这种超螺旋结构将由具有内切酶和连接酶两种活性的拓扑异构酶负责松解。DNA 拓扑异构酶有多种，主要的有拓扑异构酶Ⅰ和拓扑异构酶Ⅱ。拓扑异构酶Ⅰ能够在 DNA 双螺旋的一条链上引入一个切口，通过切口，DNA 链的断端绕过另一条链松解正超螺旋，而后两个断端再重新连接起来，该酶催化的反应不需要 ATP。拓扑异构酶Ⅰ可以同时断裂 DNA 双螺旋的两条

链，切断的 DNA 双链的两个 5′端与酶共价结合，防止造成开环，除去正超螺旋，形成负超螺旋后，再将双链连接起来，该酶催化的反应依赖 ATP。

3. 复制的终止　复制的终止是目前 DNA 复制中了解得最少的部分。DNA 的复制应该在一些特殊的终止结构区域停止。复制终止的过程包括去除 RNA 引物、冈崎片段的连接等反应，最终形成两个完整的子代 DNA 双链。

三、DNA 复制过程的共同特点

所有生物的基因组 DNA 的复制都遵循以下共同的规律和特点：

1. 半保留复制方式保证了遗传信息的忠实传递　DNA 复制时，亲代 DNA 双链解离为两条单链，各自作为模板，按碱基互补原则合成新的互补链，新合成的两个 DNA 分子和亲代 DNA 分子是完全一样的。在子代的双链 DNA 分子中一条单链来自亲代，另一条单链则是新合成的。遗传信息以这种半保留复制方式忠实地从亲代传给子代。

2. 基因组 DNA 都有固定的复制起始点　DNA 复制总是在 DNA 分子上的一个或多个位点开始。不同生物的 DNA 复制起始点的序列不尽相同，但都含有短的重复序列，这些短的重复序列可以被相关蛋白识别和结合，属于富含 AT 区域，有利于 DNA 的解链。DNA 复制时，亲代 DNA 分子在特定区域内打开双链，随之以每条单链为模板，复制生成两个子代双链 DNA 分子。

3. 复制子是基本复制单位　DNA 复制从一个复制起始点开始，直到完成一个 DNA 分子的复制或一段 DNA 的复制后停止。从一个复制起始点开始所复制的 DNA 分子或 DNA 片段是一个基本的复制单位，称为复制子（replicon）。

4. 双向复制形成复制泡和复制叉　亲代 DNA 在复制起点处解链后即形成一个复制泡，又可称为两个复制叉（replication fork），然后将沿着 DNA 分子分别向相反的方向进行双向复制（bidirectional replication）。

5. 半不连续复制方式克服了 DNA 空间结构对 DNA 新链合成的制约　DNA 双螺旋的两条链是反向平行的，两条链将同时分别作为复制的模板，而 DNA 聚合酶催化新链合成只能催化 5′→3′方向的延伸合成，且是从起点处的双向复制。如果两条链全都连续合成，意味着庞大的 DNA 分子需要先完全解链，这无疑将是一个不可克服的障碍。事实上，亲代双螺旋 DNA 链的一条链是连续被复制的，而另一条链则采用了以短片段分段的方式进行复制（图 3-1），由此避免了 DNA 空间结构的阻碍。这种复制方式被称为半不连续复制（semi-discontinuous replication）。在复制过程中，连续复制的链的前进方向始终与复制叉前进方向一致，称为前导链（leading strand）；不连续复制的链在另一条模板上的复制方向则与复制叉前进方向相反，称为随后链（lagging strand）。随后链是分段合成的，也需

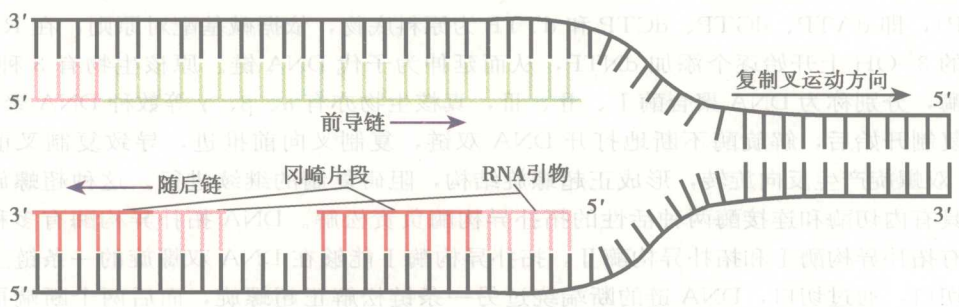


图 3-1　复制叉及半不连续复制示意图

要先合成 RNA 引物，再在引物的 3′端上延伸 DNA。这些分段合成的 DNA 片段称为冈崎片段（Okazaki fragment）。然后连接成完整的链，这样就可以保证 DNA 新链的顺利合成。

四、DNA 复制的保真性

根据核苷酸之间相互作用的测定，DNA 双螺旋碱基对之间的自由能约为 4～13 kJ/mol。这样的自由能相当于每复制 100 个碱基就有可能产生一个错配碱基，所以仅依靠 DNA 双螺旋结构的自然碱基配对原则，复制的突变率将高达 10^{-2}。而实际 DNA 复制的过程中每 10^{7}～10^{11} 个碱基对中才有可能出现一个错配的碱基。这种高度的保真性主要与 DNA 聚合酶的性质有关。

1. DNA 聚合酶识别模板并选择正确的 dNTP　DNA 聚合酶能够根据模板链上的碱基来选择正确的 dNTP 掺入到引物的 3′末端。DNA 聚合酶一般优先识别 DNA 模板和引物的 3′端，然后再识别 dNTP 与二价离子的复合物。在新的磷酸二酯键形成之前，dNTP 结合到聚合位点上，碱基对之间先形成氢键。DNA 聚合酶对正确配对的和错误配对的 dNTP 的亲和性有所不同，将其聚合至引物 3′-OH 的速度亦不同。如果引物末端的配对不正确，DNA 聚合酶能够识别出错误配对，将其排斥出聚合位点，然后以正确的 dNTP 来取代，并催化磷酸二酯键的形成。

2. DNA 聚合酶具有 3′→5′外切酶活性的校正作用　大肠杆菌 DNA 聚合酶具有 3′→5′核酸外切酶活性。DNA 聚合酶与 DNA 结合和作用可有两种状态。一是聚合状态，即引物的 3′-OH 端位于酶的聚合活性部位；二是删除状态，引物的 3′-OH 端位于酶的 3′→5′外切活性部位。酶和 DNA 的结合处于删除状态时，引物的 3′端与模板链分开，与酶的外切活性位点结合，错配的核苷酸被水解。聚合状态和删除状态之间的转换是酶与模板和引物间相互作用的结果。大肠杆菌 DNA 聚合酶Ⅰ的大片段结构域（亦称为 Klenow 片段）中有一条直径 2.2～2.4 nm 的带正电荷的深沟，DNA 进入深沟后，蛋白质结构合拢，复制时 DNA 只能在这个深沟内滑动。如果引物的 3′末端存在错配碱基，可导致引物 3′末端与 DNA 模板分开，引起 DNA 外形改变，使得 DNA 聚合酶被卡住，不能沿着 DNA 模板继续滑动，延长了引物的 3′末端与 DNA 聚合酶的 3′→5′外切活性部位之间的接触时间，将错配的核苷酸切除。DNA 聚合酶 3′→5′外切活性的即时校正作用是 DNA 复制保真性的最重要机制，可使复制的保真性提高 1～2 个数量级。

第二节　原核生物 DNA 的复制

对 DNA 复制机制的认识很多是来源于对原核生物的研究，全面了解原核生物的 DNA 复制机制，将有助于理解高等生物中的类似现象和类似规律。此外，由于细菌感染是人类疾病的一个重要致病因素，因此，对细菌基因组复制特点的理解也将有助于认识感染性疾病的某些特点。

原核生物 DNA 复制过程的主要特点是：①只有一个复制起点；②从起点向两个相反方向复制；③复制的速度快，大约每秒延伸 500 bp；④RNA 引物和冈崎片段均较长，RNA 引物的长度从十几个碱基到几十个碱基，冈崎片段约为 1 000～2 000 bp；⑤可以连续发动复制，即在上一轮复制尚未完成的情况下能发动新的复制。这里以大肠杆菌为代表，介绍原核生物 DNA 复制的机制。

一、大肠杆菌 DNA 聚合酶

大肠杆菌有三种 DNA 聚合酶。

DNA 聚合酶Ⅰ是 1 条 928 个氨基酸构成的多肽链，具有 5′→3′聚合酶活性、5′→3′外切酶活性和 3′→5′外切酶活性。每个细胞大约有 400 个酶分子，主要用于 DNA 损伤的修复。

DNA 聚合酶Ⅱ也只有一条多肽链，分子量 120 kDa，具有 5′→3′聚合酶活性和 3′→5′的外切酶活性。DNA 聚合酶Ⅱ在细菌中的具体作用尚不清楚，可能参与复制的保真控制。

DNA 聚合酶Ⅲ是 DNA 复制的主要酶，至少有 α、β、γ、δ、δ′、ε、θ、χ、Ψ 和 τ 等 10 种亚基（图 3-2）组成。这些亚基的主要作用是：①α、ε 和 θ 等 3 种亚基构成核心酶。其中 α 亚基负责催化 DNA 的合成，ε 亚基有 3′→5′的外切活性，θ 亚基则加强 ε 亚基的功能；②两个半圆形 β 亚基形成一个圆环，套在 DNA 上，形成滑卡（slide clamp）式结构。滑卡结构可以与聚合酶Ⅲ的核心酶紧密结合，随着复制叉的前进，带动核心酶沿着 DNA 滑动。即使在复制叉以每秒 1 000 bp 的速度前进时，聚合酶也不会从 DNA 上脱落；③γ 亚基与 δ、δ′、τ、χ 和 Ψ 等亚基结合而成复合体称为 γ 复合体。γ 复合体是装卡器（clamp loader），负责将封闭的 β 二聚体安装到子代 DNA 和亲代 DNA 形成的双螺旋上。其中 τ 亚基的作用是使核心酶二聚化。在复制叉上，DNA 聚合酶Ⅲ全酶形成不对称的二聚体，一侧负责合成前导链，另一侧负责合成随后链。

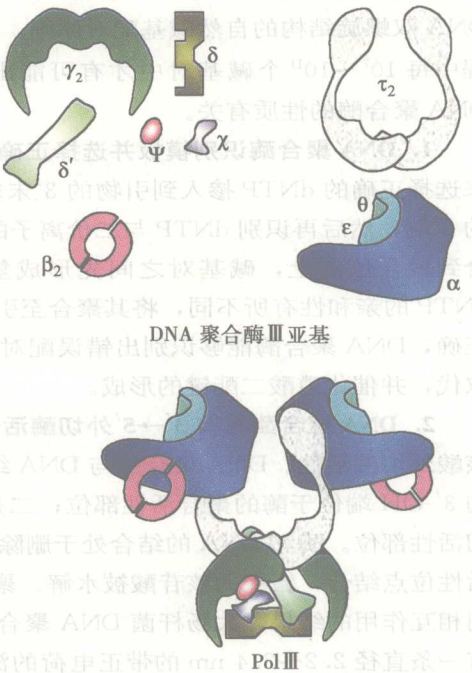


图 3-2　DNA 聚合酶Ⅲ的结构示意图

二、大肠杆菌基因组 DNA 复制的基本过程

（一）复制的起始

大肠杆菌 DNA 复制的起始包含两个步骤，一是 DNA 双链在起始部位的解链，二是形成引发体，启动 RNA 引物的合成。这一过程分别有多种蛋白分子参与。如 DnaA 蛋白、单链 DNA 结合蛋白（SSB）、DnaB 蛋白、DnaC 蛋白、DnaG 蛋白等（表 3-2）。

表 3-2　参与大肠杆菌 DNA 复制起始的主要蛋白分子

名　称	功　能	名　称	功　能
DnaA 蛋白	辨认起始点	DnaG 蛋白（引物酶）	催化 RNA 引物生成
DnaB 蛋白（解螺旋酶）	解开 DNA 双链	SSB	稳定解开的单链 DNA
DnaC 蛋白	协助解螺旋酶		

大肠杆菌的复制起点称为 *ori*C，长约 245 bp，含有两种短的重复序列（图 3-3）。其中

一种为 9 bp，另一种为 13 bp。含有 9 bp 的重复序列共有 4 个拷贝，分散在整个 *ori*C 范围内，是 DnaA 蛋白的结合位点。4 个 DnaA 蛋白与 9 bp 重复序列结合后再与其他 DnaA 蛋白结合，最后将有大约 30 个 DnaA 蛋白与 9 bp 重复序列结合。DNA 的 9 bp 重复序列区缠绕在 DnaA 蛋白上。13 bp 的重复序列共有 3 个拷贝，富含 AT 碱基，位于 *ori*C 序列的一端，串联排列。DnaA 蛋白与 9 bp 重复序列区结合后，使得 13 bp 重复序列区发生解链，单链 DNA 结合蛋白稳定局部单链结构，形成一个开放复合体。

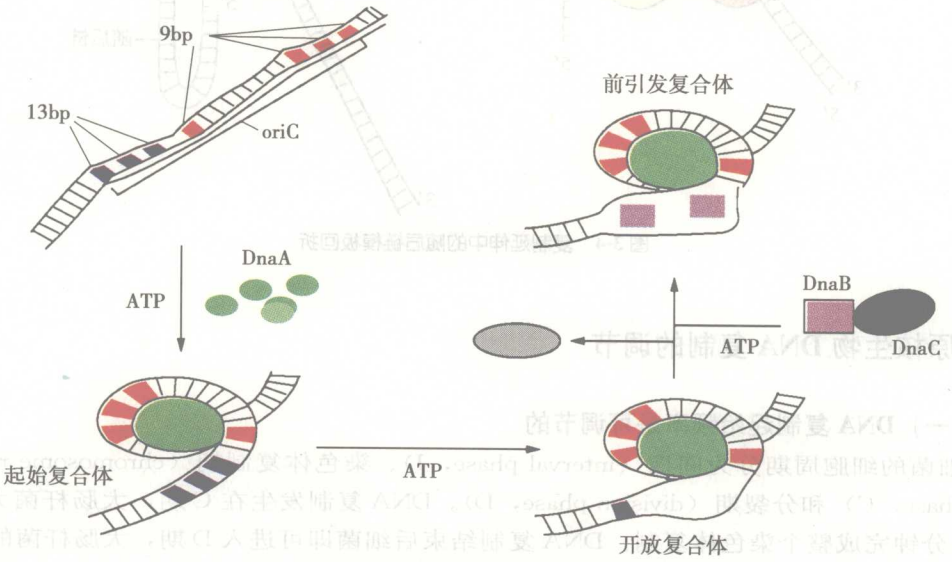


图 3-3　原核生物 DNA 复制起始示意图

复制起始点发生解链后，局部即可形成引发体。引发体是由 DnaB 蛋白、DnaC 蛋白、DnaG 蛋白和复制起始点 DNA 共同形成的复合物。引发体的蛋白质部分可在 DNA 链上移动，到达适当位置时，就可按照模板碱基序列，催化引物 RNA 的合成。

（二）复制的延伸

RNA 引物合成后，DNA 聚合酶Ⅲ即可以开始新的 DNA 链的合成。如前所述，DNA 复制中的随后链合成是不连续的。在具体合成过程中，DNA 聚合酶Ⅲ是以不对称的二聚体形式，同时合成前导链和随后链的。由于随后链合成的方向与复制叉前进的方向相反，因而随后链的模板需要在 DNA 聚合酶上折绕 180°，使得 RNA 引物的 3′端靠近 DNA 聚合酶Ⅲ的一个催化位点，才能保证随后链与前导链的合成方向都与复制叉的进行方向一致（图 3-4）。当随后链延伸至大约 1 000～2 000 bp 时，会遇到位于前方的另一个冈崎片段的 RNA 引物 5′端，此时随后链将连同模板一起脱离 DNA 聚合酶，空出的 DNA 聚合酶Ⅲ催化位点可以进入下一个有 RNA 引物的随后链模板，开始另一段随后链的合成。冈崎片段间的缺口在 DNA 连接酶催化下以 3′,5′磷酸二酯键连接起来。

（三）复制的终止

大肠杆菌基因组复制的终点是在复制起点对侧的终止区域内，终止区域中心（所谓陷阱区）两侧的终止序列可以与特异性的 Tus（terminator utilization substance，Tus）蛋白结合，使复制叉无法通过。大肠杆菌基因组的两个复制叉均终止于同一位点或其附近。复制所得到的两个环状 DNA 分子，最终仍会有 20～30 bp 互相缠绕在一起，要在拓扑异构酶Ⅱ的参与下，经过分环才能得到两个独立的环状 DNA 分子，分别被分配到两个子代细胞中去。

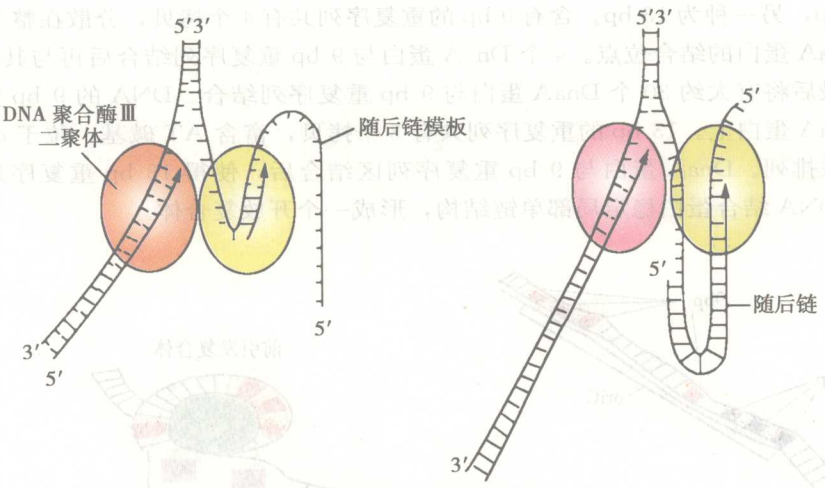


图 3-4　复制延伸中的随后链模板回折

三、原核生物 DNA 复制的调节

（一）DNA 复制起始频率是可调节的

细菌的细胞周期分为间期（interval phase，I）、染色体复制期（chromosome replication phase，C）和分裂期（division phase，D）。DNA 复制发生在 C 期，大肠杆菌大约需要 40 分钟完成整个染色体复制。DNA 复制结束后细菌即可进入 D 期，大肠杆菌的 D 期大约需 20 分钟，在此期间合成细菌分裂所需的其他细胞内成分。C 期和 D 期所需时间相对固定不变，只有 I 期的时间是可以改变的。细菌的生长和繁殖速度可因生长条件不同而改变，大肠杆菌倍增时间最短的只有 20 分钟，比 C 期和 D 期所需要的时间还要短，造成这种效应的机制是细菌的 DNA 复制的起始频率是可调节的。快速分裂的细菌复制起始频率高，在一轮复制结束前可以连续发动新的复制；而分裂缓慢的细菌复制起始频率则较低。

复制起始频率调节的分子机制尚不完全清楚，一些复制起始相关因子的浓度应该对此过程有所影响。如 DnaA 蛋白对复制起始具有正调控作用，它的浓度在某种程度上可以决定复制的起始频率。在复制起始位点的 DnaA 只有达到一定浓度才能使复制起始。DNA 复制后发生细胞分裂，复制起点局部的 DnaA 浓度降低，只有重新积累到一定浓度，才能起始下一轮复制。如果细菌生长条件好，细胞合成代谢活跃，相应的复制起始因子的浓度将很快积累到启动下一轮复制所需要的水平。

（二）复制频率与细胞分裂次数相适应

为保证每一个新的细胞都有一套完整复制的遗传信息，完成复制的 DNA 又都能够准确地分配给新的细胞，细菌细胞的基因组复制频率必须与细胞分裂周期相适应。目前认为，复制起点的甲基化修饰是控制机制之一。

大肠杆菌基因组的复制起点 *ori*C 的 13 bp 和 9 bp 两种重复序列之间有 11 个拷贝的 GATC 序列。该序列中的 A 的甲基化修饰对 DNA 复制与细胞周期相一致有重要意义。细菌中的 Dam 甲基化酶负责催化 GATC 序列中 A 的 6 位 N 的甲基化。在复制开始前，*ori*C 中 GATC 序列中的 A 必须发生甲基化，才能使复制起点活化。复制后生成的子代 DNA 分子，来自亲代单链的 A 是甲基化的，而新合成单链的 A 未被甲基化，所以子代 DNA 是半

甲基化的。一种特殊的蛋白 SeqA 能够专一地与半甲基化的 DNA 结合，并使 DNA 暂时与细胞膜相结合，膜上有甲基化的抑制因子，防止 DNA 被甲基化。当细胞分裂完成后，DNA 才与膜脱离，此时 Dam 甲基化酶才能将子代 DNA 链中的 A 甲基化，重新激活复制起点，开始下一轮复制。

第三节　真核生物染色体 DNA 的复制

真核生物基因组包括细胞核内的全部 DNA 以及线粒体的全部 DNA，因此真核生物基因组比原核生物大得多，染色体 DNA 又以核小体为单位形成了高度致密的结构。可以想象，真核生物全基因组的复制将是一个十分复杂的过程。目前对此过程的详细化学细节和调控机制仍缺乏系统了解，这里仅归纳一些目前已知的主要过程和特点。

一、真核生物染色体 DNA 与原核生物 DNA 复制的主要不同点

1. 复制时需要解开和重新组装核小体结构　真核生物的 DNA 缠绕在组蛋白八聚体上，形成核小体。复制叉经过时需要解开核小体，复制后还要重新形成核小体。真核生物 DNA 复制时，需要克服核小体结构的空间阻碍，因此，复制叉前进的速度慢。真核生物染色体 DNA 合成速度大约每秒钟 50 个核苷酸，仅为原核生物的 1/10。组蛋白的合成在细胞周期的 S 期与 DNA 复制同步进行，DNA 边合成边组装成核小体。

2. 多起点复制　真核生物染色体 DNA 分子很大，与组蛋白紧密结合，以染色质的形式存在，故真核生物 DNA 复制过程更为复杂，复制速度比原核生物慢。但是真核生物染色体 DNA 上有多个复制起始点，它们相距约 5～300 kbp，复制子小，可以在多个复制点上同时进行复制，所以真核生物 DNA 从总体上可以快速进行合成。

3. RNA 引物及冈崎片段的长度均小于原核生物　原核生物 DNA 复制过程中，RNA 引物的长度从十几个到几十个核苷酸，冈崎片段约为 1 000～2 000 个核苷酸；动物细胞中的引物约为 10 个核苷酸，冈崎片段约有 100～200 个核苷酸。真核生物中这种较短的冈崎片段可能与核小体的结构形成相适应。

4. 真核生物染色体全部复制完成以前不会发动新一轮复制　原核生物 DNA 在一轮复制结束之前可以启动下一轮复制，在快速生长的原核生物中，在起始点上可以连续开始新的 DNA 复制。而真核生物染色体 DNA 的复制则受到严格的控制。真核生物染色体 DNA 在全部复制完成以前，各个起始点上不能再开始下一轮的 DNA 复制。

二、真核生物染色体 DNA 复制过程

对于真核生物基因组 DNA 复制，研究得最为详细的是酵母细胞。本节以酵母为例介绍真核生物 DNA 复制过程中的一些特点，并简单论及不同真核生物基因组 DNA 复制的差异。

（一）真核生物 DNA 复制的 DNA 聚合酶与辅助因子

真核生物基因组复制需要数种 DNA 聚合酶及其他蛋白质分子，它们的作用和特点见表 3-3。

（二）真核生物 DNA 复制的起始和终止

真核生物 DNA 的复制从多点起始。目前对于高等动物的 DNA 复制起始位点的结构特征尚不了解，只有酵母 DNA 的复制起点是已知的。

表 3-3　真核生物染色体 DNA 复制所需要的 DNA 聚合酶及其他蛋白因子

名称	活性	作用	其他
DNA 聚合酶 α	聚合酶 引发酶	合成 8～10 个核苷酸的 RNA 引物，再合成约 30 个核苷酸的 DNA，然后改由 δ 和 ε 酶延伸 DNA 链	无 3′→5′外切酶活性
DNA 聚合酶 δ	聚合酶 3′→5′外切酶	延伸 DNA 链	需 PCNA 为辅因子
DNA 聚合酶 ε	聚合酶 3′→5′外切酶	延伸 DNA 链	不依赖 PCNA
DNA 聚合酶 β		DNA 损伤的修复	
DNA 聚合酶 γ		线粒体基因组的复制	
PCNA	DNA 聚合酶 δ 的辅助蛋白	沿 DNA 滑动，使聚合酶 δ 不会从 DNA 上滑落	
RFC		负责将 PCNA 套在 DNA 上	
RFA	单链 DNA 结合蛋白	维持 DNA 局部单链结构	
RNase H1	核酸酶	切除 RNA 引物，但在冈崎片段上保留一个核糖核苷酸	
FEN1	核酸内切酶	切除冈崎片段上保留的一个核糖核苷酸	

PCNA：proliferating cell nuclear antigen，增殖细胞核抗原
RFC：replication factor C，复制因子 C
RFA：replication factor A，复制因子 A
FEN1：flap endonuclease 1，侧翼核酸内切酶 1

1979 年发现的自主复制序列（autonomously replicating sequence，ARS）就是酵母基因组的复制起点，其核心共用序列为 A/TTTTARTTTA/T。ARS 全长不足 200 bp，由 4 个区段组成。其中的 A 和 B1 区段组成起始识别序列（origin recognition sequence，ORS），长约 40 bp，是起始识别复合物（origin-recognition complex，ORC）的结合位点。起始识别复合物由 6 个蛋白组成，以复合体的形式存在于细胞核中，专一与 A 和 B1 区段结合。区段 B2 相当大肠杆菌中的 13 bp 的重复序列区，是 DNA 双螺旋最先解链的位置。区段 B3 是 ARS 结合因子 1（ARS binding factor 1，ABF1）的结合部位。ABF1 与 B3 区段结合后，产生扭转张力，导致 B2 区段解链。复制起点双链解开后，解旋酶和其他与复制有关的酶即可与 DNA 结合，使复制叉沿 DNA 链延伸，完成复制的起始过程（图 3-5）。

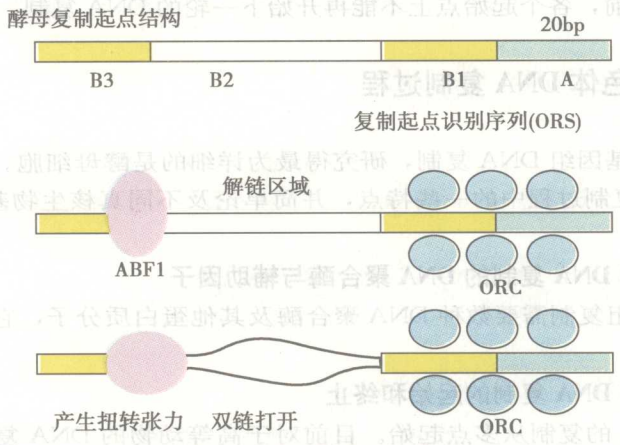


图 3-5　酵母基因组复制的起始

近年的研究还发现，在哺乳动物细胞基因组中存在一些与复制起始有关的散在序列。这些序列与酵母和其他低等真核生物的复制起点有许多相似之处。在人类 11 号染色体 β 珠蛋白基因簇一个 8 kbp 的片段中包含一个人类基因组的复制起始区。在这个复制起始区内存在一些使复制起始高频率发生的基本位点（primary site）。在基本位点的附近还同时存在一些使复制起始发生频率相对较低的次级位点（secondary site）。在人类基因组中还发现了一些基因，其编码的蛋白质与酵母的起始识别复合物中的蛋白有相似的序列，其中某些蛋白还能够替换酵母起始识别复合物中的蛋白。这些都说明哺乳动物 DNA 的复制与酵母有相似之处。酵母 DNA 复制的起始对于研究哺乳动物基因组复制的起始是一个极好的模式。

真核细胞基因 DNA 复制的起始，包括 RNA 引物的合成和新的 DNA 链合成由 DNA 聚合酶 α 完成。但是，DNA 聚合酶 α 的延伸能力较低，故在延伸阶段很快被高延伸能力的 DNA 聚合酶 δ 所取代。DNA 聚合酶 δ 取代 DNA 聚合酶 α 的过程称为聚合酶的切换（polymerase switching）。

DNA 聚合酶 δ 在复制叉上具有高延伸能力的关键原因在于它们与被称为增殖细胞核抗原（proliferation cell nuclear antigen，PCNA）的蛋白质间的结合。PCNA 由多个相同的亚基构成，并组装成“油炸圈饼”的形状，中央的空腔足以环绕 DNA 双螺旋。PCNA 沿着 DNA 滑动，但并不与 DNA 分离，并与 DNA 聚合酶 δ 紧密结合。聚合酶 δ 与 PCNA 所形成的复合体可以沿着 DNA 模板高速移动。通过保持 DNA 聚合酶 δ 与 DNA 的紧密接触，PCNA 大大增加了 DNA 聚合酶 δ 的延伸能力（图 3-6）。目前在研究中常用 PCNA 分子的含量作为反映细胞增殖状态的指标。

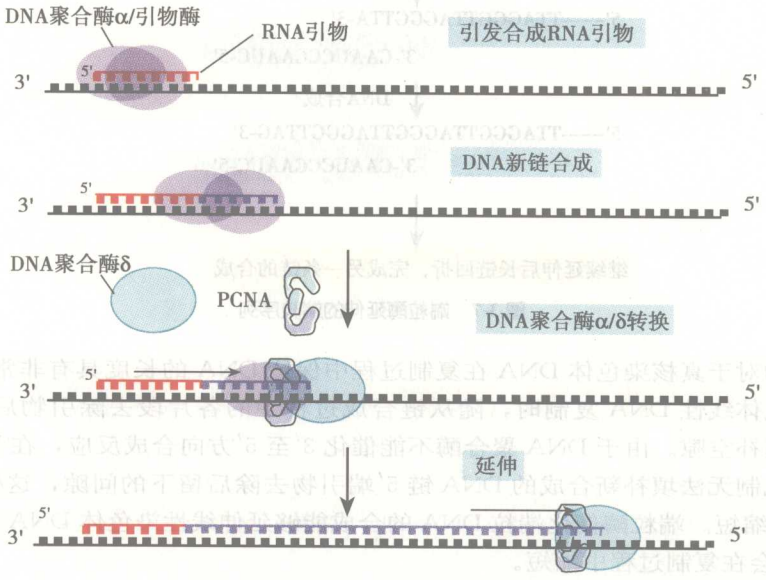


图 3-6　DNA 聚合酶 α/δ 转换在复制延伸中的作用示意图

在真核细胞中，至今还没有发现与细菌终止序列相同的序列和与 Tus 蛋白相似的蛋白，所以对真核生物复制的终止还知之甚少。

（三）端粒结构和端粒酶在真核生物基因组复制过程中的作用

真核生物基因组为线性 DNA，其末端是特殊的端粒结构。端粒（telomere）是位于真核生物染色体末端的、由 DNA 和蛋白质组成的一膨出结构。端粒 DNA 由短的高度重复

序列组成，不同生物的重复序列不同。人的端粒中，一条 DNA 链的重复序列为 TTAGGG，另一条为 CCCTAA。端粒重复序列部分可在 DNA 复制过程中以正常方式复制，部分由端粒酶来延伸。

端粒酶（telomerase）是反转录酶，由蛋白质和 RNA 共同组成，能以自身的 RNA 为模板反复延伸端粒的重复序列。端粒酶 RNA 中作为模板的序列与端粒重复序列相互补，长度为端粒重复序列的 1.5 倍。人的端粒酶中的 RNA 长度为 450 个核苷酸，5′端作为模板的核心区序列为 5′-UAACCCUAA-3′，与端粒 DNA 重复序列 5′-TTAGGG-3′呈反向互补。如图 3-7 所示，端粒的重复序列 TTAGGG 先与端粒酶 RNA 模板序列 3′-AAUCCCAAU-5′中的 AAUCCC 配对，模板中没有配对的 AAU 指导合成新的 TTA。然后，端粒酶移位，使模板 3′端的 AAU 与新合成的 TTA 配对，模板 CCCAAU 指导合成新的 GGGTTA，端粒重复序列增加了一个拷贝。由于端粒酶不断移位，GGGTTA 重复序列不断得以延伸。端粒酶将端粒重复序列延伸到一定的长度后，端粒重复序列靠 G-G 非标准配对回折，提供 3′-OH，以已合成的端粒重复序列为模板合成另一条链上的端粒重复序列。

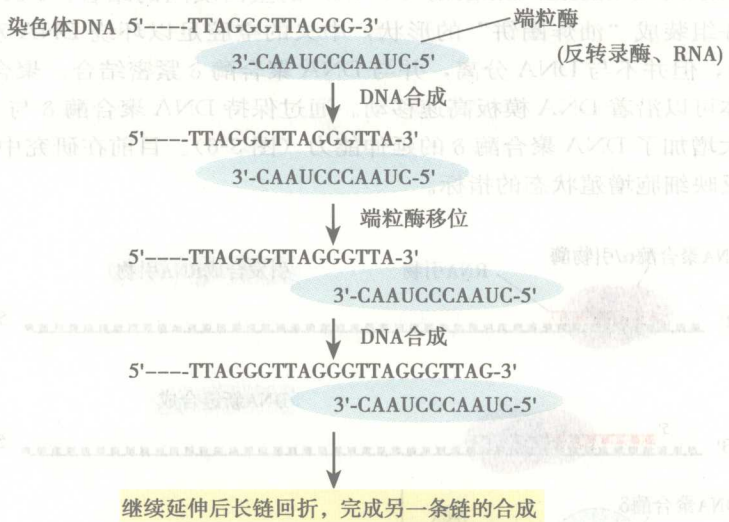


图 3-7　端粒酶延伸的端粒序列

端粒结构对于真核染色体 DNA 在复制过程中保持 DNA 的长度具有非常重要的作用。真核生物染色体线性 DNA 复制时，随从链合成过程中的各片段去除引物后，是由 DNA 聚合酶 δ 来填补空隙。由于 DNA 聚合酶不能催化 3′至 5′方向合成反应，在 DNA 的末端，靠常规复制机制无法填补新合成的 DNA 链 5′端引物去除后留下的间隙，这样会造成染色体 DNA 逐渐缩短。端粒酶催化端粒 DNA 的合成能够延伸线性染色体 DNA 的末端，保证 DNA 分子不会在复制过程中缩短。

端粒 DNA 的长度由端粒结合蛋白（telomere binding protein，TBP）调控。一定量的 TBP 与端粒结合后可以抑制端粒酶的活性。适度的端粒酶活性对于细胞的正常增殖非常重要。端粒重复序列长度随细胞分裂次数的增加和年龄的增长而缩短。端粒 DNA 长度逐渐变短甚至消失，可导致染色体稳定性下降，引起细胞衰老。反过来，如果因某些原因引起端粒酶活性持续增强，端粒重复序列不断延长，细胞会不停地分裂，有可能导致癌变。

（四）复制过程中核小体的装配

真核生物的 DNA 缠绕在组蛋白八聚体上，形成核小体。复制叉经过时需要解开核小

体，复制后还要重新形成核小体。

组装新的核小体的组蛋白的合成发生在细胞的S期，与DNA的复制同步进行。组蛋白基因的转录起始于G1期晚期，在随后的S期继续进行。在细胞从G1期向S期过渡的过程中，有刺激组蛋白基因转录的调节蛋白产生或被激活。能和组蛋白前体mRNA的3′端碱基互补、参与3′端多余序列剪接加工的U7 snRNA片段也在S期出现。当S期开始时，成熟的组蛋白mRNA就出现在细胞质中并立即翻译成组蛋白，运回细胞核与DNA组装成新的核小体。在S期，组蛋白mRNA的半衰期约为45～60分钟，在整个S期都保持高水平地合成组蛋白。当DNA复制完成后，组蛋白基因转录关闭，成熟的mRNA停止进入细胞质，mRNA的半衰期也迅速降到10～15分钟。组蛋白基因转录的停止和细胞质中组蛋白mRNA的迅速降解，基本上终止了组蛋白的合成。保持组蛋白和DNA量比例的稳定。

三、真核生物DNA复制的调控

真核生物DNA复制的调控分为三个层次：细胞周期水平的调控、染色体复制水平的调控和复制子水平的调控。最关键的调控是与细胞周期相同步的调控。

（一）细胞周期进程中的调控

在细胞周期进行中，基因组的全部信息必须被精确地复制并均等地分配到两个子代细胞中，且复制的进程与细胞周期的进程相协调。真核生物细胞周期的S期是高等真核生物细胞染色体的复制期。DNA复制从上千个复制起始点启动，进行双向复制。在S期，有的起始点开放得早，有的起始点开放得晚，但在整个S期，所有的起始点只开放一次，保证每个周期所有的染色体都得到复制，而且只复制一次。

细胞周期蛋白D的水平在G1后期升高，激活S期的CDK（cyclin-dependent kinase，CDK）。复制许可因子（replication licensing factor）是CDK的底物，为发动DNA复制所必需。复制许可因子一般不能通过核膜进入核内，但是在有丝分裂的末期、核膜重组之前可以进入细胞核，与DNA的复制起点结合。等待被刺激进入S期的CDK激活，启动复制。一旦复制启动，复制许可因子即失去活性或被降解。在细胞周期的其他时间内，新的复制许可因子不能进入细胞核内，保证在一个细胞周期内只能进行一次基因组的复制。

目前对于真核生物基因组复制与细胞周期协调的调控机制了解得还不多，主要是在酵母等比较简单的真核生物研究中取得了一些进展。酵母基因组由17条染色体组成，各染色体又都有多个复制子。各个复制子按照专一的时间顺序活化，在S期的不同阶段开始复制。酵母基因组的复制受到蛋白因子和DNA特定序列两种因素调控，通过两者的相互作用调控基因复制的进程。真核生物基因组复制调控过程中许多细节目前所知甚少。

（二）染色体水平的调控

不同染色体或同一染色体的不同部位的复制子，按先后顺序在S期的不同时间起始复制，是什么因素控制着不同复制子的起始顺序，目前尚不清楚。

（三）复制子水平的调控

真核生物复制的起始，包括复制起始复合物的形成和RNA引物的合成等多个步骤，有许多蛋白和酶参与其中，这些步骤都是复制起始的关键和调控点。

四、线粒体DNA的复制

动物线粒体和植物叶绿体环状双链DNA的复制常采用D环复制模式。复制开始时，

双链环状 DNA 的特殊位点先解链形成一个复制泡。复制泡中的一条亲代 DNA 链作为模板合成前导链。前导链与模板链形成局部双链，将另一条亲本链置换出去。这种由一条单链和一条双链组成的泡状结构称为 D 环（displacement loop）。前导链继续合成，使另一条亲本链上的复制起点出现在单链区，随后链即开始复制。亲代双链 DNA 的两条链各有自己的复制起点。两个复制起点的激活有先有后，两条链的复制也不是同时结束（图 3-8）。因此，D 环复制的特点是两条链的复制是不同步的。

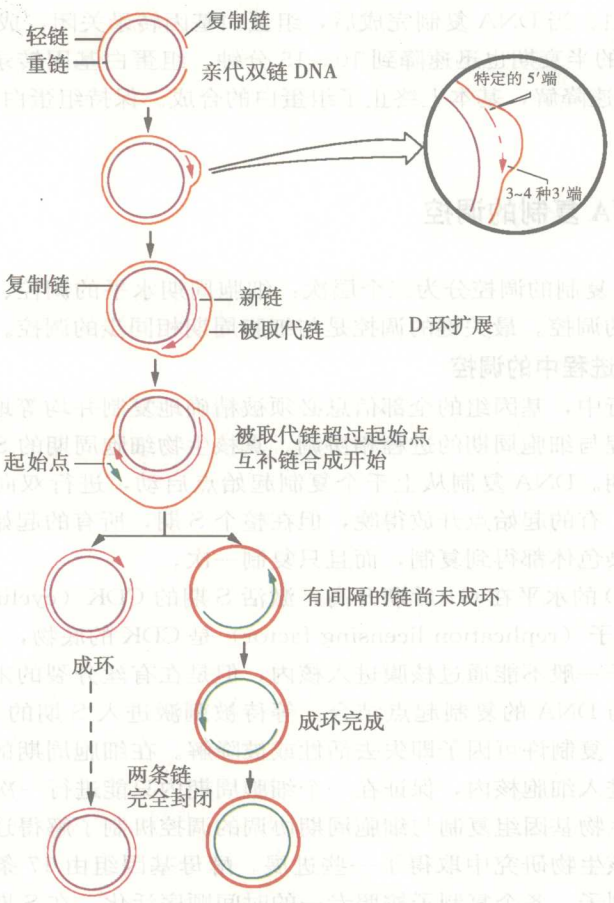


图 3-8　D 环复制模式

线粒体 DNA 上的 D 环区位于 $tRNA^{Pro}$ 和 $tRNA^{Phe}$ 基因之间，包含 3 个保守序列节段、终止结合序列（termination associated sequences，TAS）和起始点等序列。线粒体 DNA 复制由 DNA 聚合酶 γ 负责。线粒体 DNA 合成速度相当缓慢，每秒约 10 个核苷酸，整个复制过程需要 1 小时。新合成的线粒体 DNA 是松弛型的，需要 40 分钟将其变成超螺旋型。

线粒体 DNA 控制区调控整个线粒体 DNA 的复制和转录，该区的碱基突变可能引起线粒体 DNA 复制和转录速度加快，造成突变的线粒体 DNA 相对增多，经过长期的累积可能导致肿瘤的发生。此外，线粒体基因组的缺陷和衰老、肿瘤、神经退行性疾病、肌病、心肌病以及糖尿病的关系密切。

第四节　某些病毒基因组的复制

病毒基因组的复制是以病毒基因组（DNA 或者 RNA）为模板，借 DNA 聚合酶或 RNA 聚合酶以及其他因子，经过复杂的生化合成过程，复制出病毒的基因组。一些病毒通过在宿主细胞中大量复制而繁殖，进而对宿主产生致病作用。组成病毒基因组的 DNA 和 RNA 可以是单链结构，也可以是双链结构。病毒基因组的多种类型决定了其基因组复制的复杂性（尤其是 RNA 病毒）。DNA 病毒除痘类病毒外都在核内进行基因组复制，RNA 病毒（除反转录病毒外）的基因组复制全部在细胞质中进行。人和动物的 RNA 病毒多为单链 RNA 病毒，绝大多数在细胞质中进行生物合成。

不同的病毒 DNA 的复制方式不同，其共同特点一是都需要宿主细胞提供一套完成复制所需要的蛋白分子，二是 DNA 合成起始的引物形式多样。

一、腺病毒 DNA 的复制

腺病毒基因组是双链线性 DNA，两端有反向末端重复序列（inverted terminal repeat，ITR）。腺病毒基因组的复制至少需要 6 种蛋白，其中 DNA 结合蛋白（DNA binding protein，DBP）、前末端蛋白（pre-terminal protein，PTP）和 DNA 聚合酶等 3 种分子由病毒基因编码；另外三种蛋白，NF Ⅰ（nuclear factor Ⅰ，核因子Ⅰ）、NF Ⅱ、NF Ⅲ等 3 种蛋白分子由宿主细胞提供。NF Ⅰ、NF Ⅲ同时参与转录和复制过程，说明腺病毒的转录和复制机制有相同之处。DBP 的主要作用是保护单链 DNA，DNA 聚合酶催化 dCMP 上的 5′磷酸基团和 PTP 中丝氨酸的-OH 之间形成磷酸二酯键，dCMP 与模板 DNA 的 3′端配对，作为 DNA 合成的引物。除 NF Ⅰ、NF Ⅲ之外，复制还需要 NF Ⅱ。

腺病毒 DNA 开始复制时，NF Ⅰ、NF Ⅲ、DBP 首先结合到复制起点，然后 PTP 和聚合酶形成的复合物与 DNA 结合，催化腺病毒 DNA 末端解链，以 PTP 上的 dCMP-OH 为引物起始 DNA 链的复制。在 DNA 延伸的过程中，聚合酶与 PTP 解离后结合到 DBP 分子上，继续延伸 DNA 链。

腺病毒 DNA 复制最重要的特点就是由前末端蛋白 Ser 的-OH 提供引物，使 dCMP 和模板的 3′末端配对，保证线性基因组复制后 DNA 分子的 5′端不缩短。除腺病毒外，已知细菌病毒 Φ29 和 PRO1 也采用末端蛋白作为引物。

二、单链 RNA 病毒基因组的复制

单链 RNA 病毒根据 RNA 是否具有 mRNA 功能分为正单链和负单链 RNA 病毒。烟草花叶病毒、脊髓灰质炎病毒等都是正链单链 RNA 病毒，其基因组较小，只编码 A 蛋白、外壳蛋白和复制酶等。复制正链时，只需要复制酶。以正链为模板；复制负链时，则需要复制酶和宿主核糖体蛋白 S1、延伸因子 EF-Tu 和延伸因子 EF-Ts 结合成全酶。由于复制酶全酶对负链的亲和力大于对正链的亲和力，得到的正链总比负链多。正、负链只在复制酶工作的位置上结成双链。组装病毒颗粒时只包装正链，不包装负链。

RNA 复制酶催化合成互补链 RNA 的反应机制与其他聚合酶催化核酸合成反应相似。RNA 复制酶缺乏校对功能，因此 RNA 复制的错误率较高，这与 DNA 指导的 RNA 聚合反应情况相类似。RNA 复制酶只有特异地对病毒 RNA 起作用，一般不能对宿主细胞 RNA 进行复制。

水疱性口膜炎病毒（VSV）、狂犬病毒、流感病毒等都是负链单链 RNA 病毒，基因组 RNA 不能作为 mRNA。VSV 病毒复制时要得到全长的正链和负链 RNA，转录时则要得到 5 种 5′端起点不同、都有 Poly（A）尾巴的 mRNA。

流感病毒基因组由 8 条 RNA 组成。这 8 条 RNA 各自独立，长短不等，编码的蛋白参与基因组的复制和转录。流感病毒基因组转录时，要从宿主细胞窃取新生的、已经加帽的 mRNA 作引物，转录沿模板的 3′→5′的方向进行，接近模板的 5′端时会遇到 5～7 个成串的 U，转录酶在 mRNA 上加 Poly（A）尾。复制和转录用同一套酶，但要转录出全长互补的 RNA 和负链 RNA，则不能用 mRNA 作模板。组装子代病毒体时，8 条 RNA 链中各取一条，调控机制不明。

三、反转录病毒基因组的复制

反转录病毒（retrovirus）基因组 RNA 在反转录酶作用下通过 cDNA 中间体进行复制。反转录酶的作用是以 RNA 为模板指导三磷酸脱氧核苷合成 cDNA。反转录病毒是正链 RNA 病毒，其基因组 RNA 相当于 mRNA，但入侵细胞后并不编码蛋白质，而是经病毒的反转录酶将它反转录成 DNA，整合到宿主细胞的染色体上，称为原病毒。然后，再以原病毒 DNA 为模板，由宿主细胞的转录酶转录出 RNA。这种 RNA 既可以编码蛋白质，也可作为基因组。反转录病毒基因组的结构特点是两端有长末端重复序列（long terminal repeat，LTR），在 LTR 中有转录和复制的起始信号及引物的结合位点（primer binding site，PBS），LTR 中的同向重复序列对于复制起始极为重要，tRNA 引物的结合位点就在同向重复序列中。

病毒侵入细胞后，首先将 RNA 基因组反转录成 DNA。反转录酶与 DNA 聚合酶一样，不能直接合成 DNA，必须有 RNA 作为引物。反转录病毒反转录时借用宿主细胞的 tRNA 作为引物。tRNA 的 3′端和 PBS 互补。不同病毒借用的 tRNA 不同。

引物 tRNA 与病毒基因组 5′的 PBS 形成碱基配对，在 tRNA 的 3′端先添加 NMP（一磷酸核苷），随后改为添加 dNMP（脱氧一磷酸核苷），进行反转录，直到基因组的 5′端。引物 tRNA 连同反转录生成的 DLA 与病毒基因组的 5′端脱离，再与 3′端的 PBS 配对（和 5′端同向重复序列），从基因组的 3′端向 5′端反转录出整个 DNA 负链。当负链 DNA 延伸到一定长度时，内切酶在正链 RNA 的 5′切断，以产生的 3′-OH 为引物，DNA 负链为模板，合成正链 DNA，得到的双链 DNA 整合到细胞基因组，随细胞基因组转录出 RNA。这些 RNA 既是 mRNA，又可作为病毒新的基因组。

四、乙型肝炎病毒基因组的复制

乙型肝炎基因组是双链 DNA。负链 DNA 为长链，呈环状但不闭合，因其 5′端有蛋白分子结合，因而妨碍 3′,5′磷酸二酯键的形成。正链是短链，长短不一，3′端和 DNA 聚合酶结合，需要时可延伸。正链作为桥连封堵负链 DNA 的缺口，使负链 DNA 形成环状。

乙型肝炎病毒基因组复制的特点是复制过程中要经过 RNA 转录阶段。病毒侵入细胞后先将正链 DNA 延伸到全长，成为双链 DNA，再依次以负链 DNA 为模板，转录出正链 RNA；以正链 RNA 为模板，反转录出负链 DNA；以负链 DNA 为模板，合成正链 DNA；正链 DNA 将负链 DNA 桥连成环。

（高　旭）

第四章　DNA 损伤与修复

关键知识点

- DNA 损伤是指各种体内外因素所导致的 DNA 组成和结构变化。DNA 损伤表现为碱基和糖基的破坏、碱基错配、DNA 单链或双链断裂、DNA 链的共价交联等多种类型。
- 细胞内在因素如 DNA 复制中的错配、DNA 结构本身的不稳定性、机体代谢中产生的有害物质等均可诱发 DNA 的“自发”损伤。外环境中的物理因素（电离辐射、紫外线）、化学因素（自由基、碱基类似物、碱基修饰剂、嵌入性染料）和生物因素等均可损伤 DNA。
- DNA 损伤具有双重生物学效应。各种因素诱发的 DNA 结构改变是生物进化的基础；但是 DNA 的损伤也可使细胞功能出现障碍，与多种疾病有关。
- 生物存在 DNA 损伤修复机制，可以纠正碱基的错配，清除 DNA 链上的损伤，恢复 DNA 正常结构。这一机制对于维持 DNA 结构的完整性与稳定性，保证生命延续和物种稳定至关重要。细胞有直接修复、切除修复、重组修复和损伤跨越等多种修复 DNA 损伤的途径。一种 DNA 损伤可通过多种途径来修复，一种修复途径也可参与多种 DNA 损伤的处理。DNA 损伤修复的缺陷与肿瘤、衰老、免疫缺陷等多种疾病的发生有着密切关联。

关键词： DNA 损伤；无嘌呤或嘧啶位点；光复活酶；着色性干皮病（XP）；DNA 损伤修复；碱基切除修复；核苷酸切除修复；同源重组修复

遗传物质 DNA 的遗传保守性是维持物种相对稳定的主要因素。然而，在长期的生命过程中，生物体时刻受到来自内、外环境中各种因素的影响，DNA 的改变不可避免。各种体内外因素所导致的 DNA 组成和结构变化称为 DNA 损伤（DNA damage）。DNA 损伤可有两种后果：一是导致 DNA 失去作为复制和转录的模板的功能，导致细胞的死亡；二是 DNA 的结构发生永久性的改变即突变。

在长期的进化中，无论低等生物还是高等生物都形成了自己的 DNA 修复系统，能随时纠正和修复 DNA 损伤，恢复 DNA 的正常结构，保持细胞的正常功能。DNA 损伤的同时即伴有 DNA 修复系统的启动。受损细胞的转归在很大程度上取决于 DNA 的修复效果，如能正确修复，细胞 DNA 结构恢复正常，细胞得以维持正常状态；如损伤严重，DNA 不能有效修复，则通过凋亡来清除这些细胞，从而降低了 DNA 损伤对生物体遗传信息稳定性的影响；当 DNA 发生不完全修复时，DNA 发生突变，染色体发生畸变，可诱导细胞出现功能改变、衰老、恶性转化等病理变化。但是如果遗传物质具有绝对的稳定性，那么生物将会失去进化的机制，就不会呈现大千世界、万物生辉的自然界。因此，生物多样性依赖于 DNA 突变与 DNA 修复之间的良好平衡。

第一节　DNA 损伤

DNA 损伤的诱发因素众多。一般可分为环境因素（外源）和生理因素（内源）。前者包括辐射、化学毒物、药物、病毒感染、植物和微生物代谢产物等。后者是指机体代谢过程中产生的有毒物质、DNA 复制过程中发生的错配和 DNA 本身的热不稳定性因素等，这些因素可诱发 DNA 的“自发”损伤。值得注意的是环境因素和生理因素不能截然分开，

多种环境因素都可通过与生物体的相互作用产生有毒物质，进而导致 DNA 损伤。不同因素引起的损伤机制亦不相同。

一、导致 DNA 损伤的多种因素及其机制

（一）生物个体内在因素

1. DNA 复制错误　在 DNA 进行复制的过程中，4 种 dNTP 浓度的不平衡、碱基的异构互变等都可以引起碱基的错配，即产生非 Watson-Crick 碱基对。尽管绝大多数错配的碱基会在复制中被 DNA 聚合酶的即时校读功能所纠正，但不可避免仍有少数被保留下来。DNA 复制过程的错误率约为 10^{-10}。

此外，复制错误还可以表现为片段性的缺失与插入，造成 DNA 序列的改变。特别是短片段重复序列，因为其重复片段在人类和其他真核细胞染色体上的分布广泛，导致 DNA 复制系统工作时可能出现“打滑”现象，使得新生成 DNA 上的重复序列拷贝数目减少或增加，微卫星 DNA 重复片段在长度方面表现出高度的多态性，在可遗传性突变疾病研究上有重大价值。人类的三核苷酸重复序列扩增（triplet-repeat expansion）正是由于重复序列 CAG 的拷贝数在世代相传的过程中发生了扩增，造成基因编码的蛋白产物功能障碍，致使相应的疾病在患病个体的后代中进行性加重。亨廷顿病、脆性 X 综合征、肌强直性营养不良等神经退行性疾病均属于此类疾病。

2. DNA 自身的不稳定性　DNA 结构本身的不稳定性是 DNA 自发性损伤中最频繁和最重要因素。当 DNA 受热或是所处的 pH 值发生改变时，DNA 分子上连接碱基和核糖之间的糖苷键自发地发生水解，导致碱基的丢失或脱落，以脱嘌呤最普遍。含有氨基的碱基还可以自发地发生脱氨基反应，从而转换为另一种碱基即碱基的转换，如 C 经脱氨基反应转变为 U，A 脱氨基转变为次黄嘌呤（I）等。

3. 机体代谢过程中产生的活性氧　机体代谢过程中产生的活性氧（reactive oxygen species，ROS）可以直接作用于碱基，修饰鸟嘌呤产生 8-羟基脱氧鸟嘌呤。

（二）环境因素

最常见的导致 DNA 损伤的环境因素包括物理因素、化学因素和生物因素，其导致 DNA 损伤的机制各有特点。

1. 物理因素　物理因素中最常见的是辐射致 DNA 损伤。根据作用原理不同，通常将辐射分为电离辐射和非电离辐射。α 粒子、β 粒子、X 射线、γ 射线等，能直接或间接引起被穿透物质产生电离，属电离辐射；紫外线和能量低于紫外线的所有电磁辐射均属非电离辐射。

（1）电离辐射导致的损伤：电离辐射既可直接作用于生物大分子导致结构的破坏，如化学键的断裂等；又可通过在细胞内激发形成自由基等反应分子而起间接作用。这些作用最终导致 DNA 分子发生碱基氧化修饰、碱基环结构的破坏与脱落、DNA 链交联与断裂等多种变化。

（2）紫外线照射导致的损伤：紫外线（ultraviolet，UV）属非电离辐射。按波长的不同，紫外线可分为 UVA（400～320 nm）、UVB（320～290 nm）和 UVC（290～100 nm）3 种。UVA 的能量较低，一般不造成 DNA 等大分子的损伤。大气臭氧层可吸收大部分 320 nm 以下的紫外线，一般不会造成地球上生物的损害。但近年由于环境污染，臭氧层的破坏日趋严重，UV 对生物的影响越来越为公众所关注。260 nm 左右的紫外线，其波长正好在 DNA 和蛋白质的吸收峰附近，容易导致 DNA 等大分子的损伤。

紫外线高能量的吸收可使 DNA 分子中同一条链两相邻胸腺嘧啶碱基（T）间以共价

键连接形成胸腺嘧啶二聚体结构（TT），或称为环丁烷型嘧啶二聚体（图 4-1）。紫外线也可导致其他嘧啶间形成类似的二聚体如 CT、CC。二聚体的形成可使 DNA 产生弯曲和扭结，影响 DNA 双螺旋结构，使复制和转录受阻。另外，紫外线还会导致 DNA 链交联和链断裂等损伤。

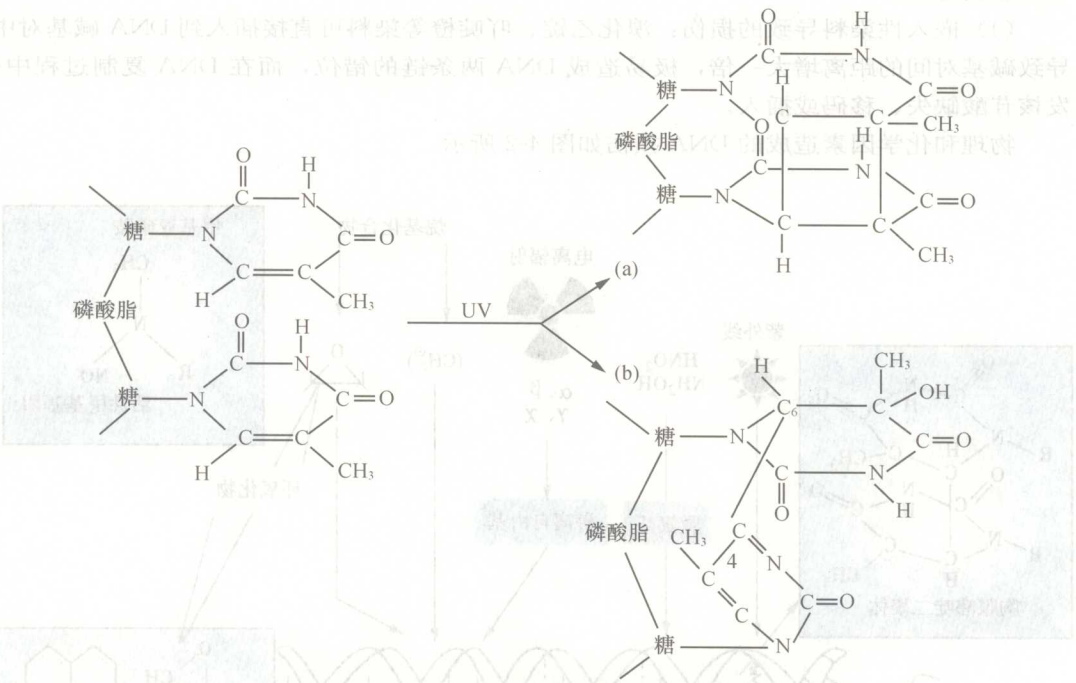


图 4-1　胸腺嘧啶二聚体的形成

2. 化学因素　能引起 DNA 损伤的化学因素种类繁多，按其作用原理可分为自由基、碱基类似物、碱基修饰物和嵌入染料。值得注意的是，许多肿瘤化疗药物是通过诱导 DNA 损伤包括碱基改变、单链或双链 DNA 断裂等来阻断 DNA 的复制和转录，从而抑制了肿瘤细胞的增殖。因此，对 DNA 损伤及后继的细胞死亡机制的认识，有助于对肿瘤化疗药物的改进。

（1）自由基导致的损伤：自由基是指能够独立存在，核外带有未配对电子的原子或分子。自由基的化学性质异常活跃，可引起多种化学反应，对细胞功能产生影响。自由基的产生可以是外界因素与体内物质相互作用的结果，如电离辐射产生的羟自由基（·OH）和氢自由基（·H）。·OH 具有极强的氧化性质，而·H 则具有极强的还原性质。生物体内代谢过程产生的活性氧等。这些自由基可通过化合反应、抽氢作用、加成反应等与 DNA 分子发生相互作用，进而导致碱基、核糖、磷酸基的损伤，引起 DNA 的结构和功能的异常。

（2）碱基的某些类似物导致的损伤：通常是人工合成的一类与 DNA 正常碱基结构类似的化合物，用作促突变剂或抗癌药物。在 DNA 复制时，因其结构相似取代正常碱基掺入并与互补链上碱基配对，从而引起碱基对的置换。5-溴尿嘧啶（BU）是胸腺嘧啶的类似物，有酮式和烯醇式结构，前者与腺嘌呤配对，后者与鸟嘌呤配对。可导致 DNA 中 AT 配对转变为 GC 配对，GC 配对转变为 AT 配对。

（3）碱基修饰剂、烷化剂导致的损伤：这是一类通过对 DNA 链中碱基的某些基团的修饰，改变其配对性质，进而改变 DNA 结构的化合物。例如亚硝酸能脱去碱基上的氨基，腺嘌呤脱氨后成为次黄嘌呤，不能与原来的胸腺嘧啶配对，而与胞嘧啶配对；胞嘧啶脱氨

为尿嘧啶，转而与腺嘌呤配对，从而改变碱基序列。此外众多的烷化剂如氮芥、硫芥、二乙基亚硝胺等可导致 DNA 碱基上的氮原子烷基化，引起分子电荷的变化，从而改变碱基配对，或烷基化鸟嘌呤脱落形成无碱基位点，或引起 DNA 链中的鸟嘌呤连接成二聚体，或导致 DNA 链交联与断裂。这些变化都可以引起 DNA 序列或结构的异常，并可阻止正常的修复过程。

（4）嵌入性染料导致的损伤：溴化乙锭、吖啶橙等染料可直接插入到 DNA 碱基对中，导致碱基对间的距离增大一倍，极易造成 DNA 两条链的错位，而在 DNA 复制过程中引发核苷酸缺失、移码或插入。

物理和化学因素造成的 DNA 损伤如图 4-2 所示。

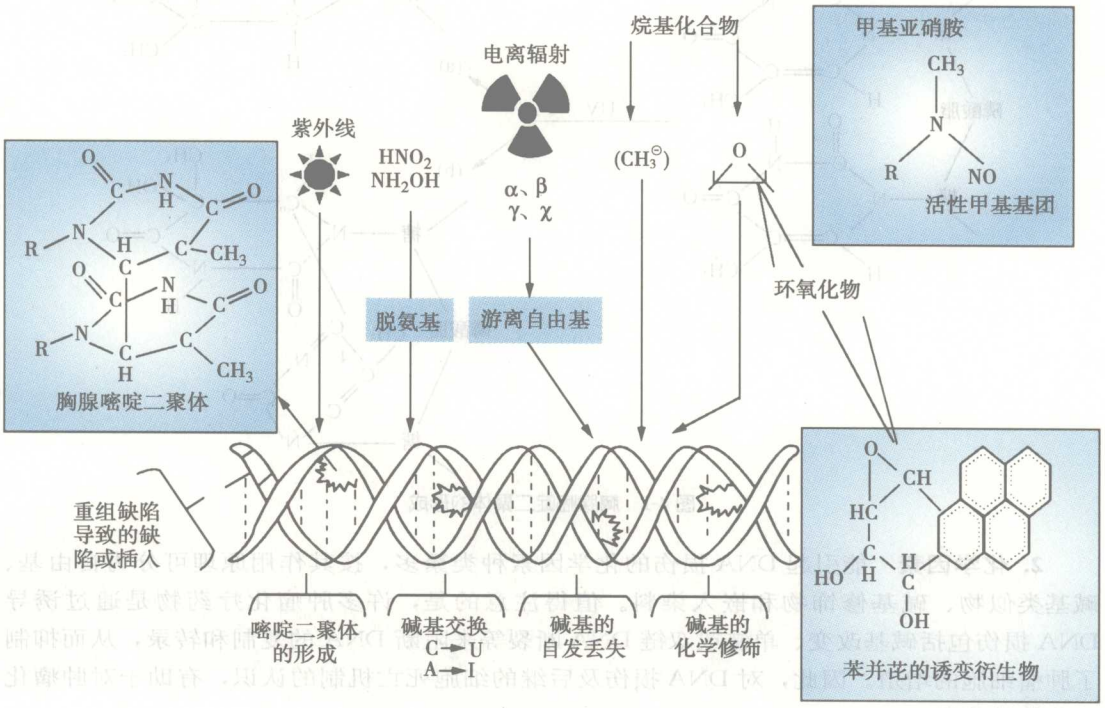


图 4-2　物理和化学因素对 DNA 的损伤

3. 生物因素　生物因素主要指病毒。如麻疹病毒、风疹病毒、疱疹病毒等，它们产生的毒素和代谢产物，如黄曲霉素等都有诱变作用。

二、DNA 损伤以多种类型存在

DNA 分子中的碱基、核糖和磷酸二酯键等都是致 DNA 损伤因素作用的靶点。根据 DNA 分子结构改变的不同分为碱基脱落、碱基破坏、嘧啶二聚体形成、DNA 单链和双链断裂、DNA 交联等多种类型。

1. 碱基和糖基的破坏　化学毒物可通过对碱基的某些基团进行修饰而改变碱基的性质，如亚硝酸能导致碱基脱氨；嘧啶碱基在羟自由基的攻击下易发生加成、抽氢等反应，导致碱基环的破裂；具有氧化活性的物质可能造成 DNA 中嘌呤和嘧啶碱基的氧化修饰，形成 8-羟基脱氧鸟苷或 6-甲基尿嘧啶氧化代谢产物；紫外线作用于 DNA 分子可形成嘧啶二聚体；糖基的碳原子和羟基上的氢也可能与自由基反应。由于碱基或糖基的损伤，在

DNA 链上可能形成一些不稳定点，最终可导致 DNA 链的断裂。

2. 碱基之间发生错配　如前所述，碱基类似物的掺入、碱基修饰剂的作用可改变碱基的性质，导致 DNA 序列中的错误配对。在正常 DNA 复制过程中也存在一定比例自发的碱基错配发生。最常见的是组成 RNA 的尿嘧啶替代胸腺嘧啶掺入 DNA 分子中。

3. DNA 链发生断裂　DNA 链断裂是电离辐射致 DNA 损伤的主要形式，某些化学毒剂也可导致这种损伤。磷酸二酯键的断裂、脱氧戊糖的破坏、碱基的破坏和脱落都是引起 DNA 断裂的原因。碱基或糖基的损伤可引起 DNA 双螺旋局部变性，形成酶敏感位点，特异的核酸内切酶能识别并切割这样的损伤而造成链断裂。DNA 链上损伤的碱基也可以被另一种特异的 DNA-糖基化酶除去，形成无嘌呤或无嘧啶位点（apurinic/apyrimidinic site，AP 位点）或称无碱基位点，这些位点在内切酶等的作用下可形成链断裂。DNA 断裂可以发生在单链或双链上，单链断裂能迅速在细胞中以另一链为模板重新合成而完整修复，但双链断裂在原位修复的几率很小，需依赖重组修复，这种修复导致染色体畸变的可能性很大。因此，一般认为双链断裂的 DNA 损伤与细胞的致死效应有直接的联系。

4. DNA 链的共价交联　DNA 损伤产生多种的交联形式，DNA 双螺旋链中的一条链上的碱基与另一条链上的碱基以共价键结合，称为链间交联（DNA interstrand cross-linking）。DNA 分子中同一条链中的两个碱基以共价键结合，称为 DNA 链内交联（DNA intrastrand cross-linking）。DNA 分子还可与蛋白质以共价键结合，称为 DNA-蛋白质交联（DNA protein cross-linking）。紫外线照射后形成的嘧啶二聚体就是 DNA 链内交联的典型例子。

以上对各种类型 DNA 损伤进行了阐述，实际上 DNA 损伤是相当复杂的。当 DNA 受到严重损伤时，在局部范围发生的损伤常不止一种，而是多种类型的损伤复合存在。最常见的是碱基损伤、糖基破坏和链断裂，这种损伤部位被称为局部多样损伤部位。

上述 DNA 损伤可导致 DNA 模板发生碱基置换、插入、缺失、链的断裂等变化，并可能影响到染色体的高级结构。就碱基置换来讲，DNA 链中一种嘌呤被另一种嘌呤取代，或嘧啶被另一种嘧啶所取代，称为转换；嘌呤被嘧啶取代或反之，称为颠换。转换和颠换在 DNA 复制时可引起碱基错配，导致基因突变。碱基的插入和缺失可引起移码突变。DNA 断裂阻止了 RNA 合成过程中链的延伸。DNA 损伤引起的染色体结构变化也可造成转录和翻译的异常。所有这些变化可造成某种（些）基因信息发生丢失或异常，进而导致其表达产物量和质的变化，对细胞的功能造成不同程度的影响。

第二节　DNA 损伤的修复

生物体在长期的生命活动中不可避免的都会发生 DNA 损伤。DNA 损伤的后果取决于 DNA 损伤程度和细胞的 DNA 损伤的修复能力。DNA 损伤修复是指纠正错配的碱基、清除 DNA 链上的损伤、恢复 DNA 正常结构的过程。DNA 修复是机体维持 DNA 结构的完整性与稳定性，保证生命延续和物种稳定的重要环节。

细胞内存在多种修复 DNA 损伤的途径。常见的 DNA 修复方式包括：直接修复、切除修复、重组修复和损伤跨越等（表 4-1）。值得注意的是，一种 DNA 损伤可通过多种途径来修复，一种修复途径也可参与多种 DNA 损伤的处理。

一、某些 DNA 损伤可以直接修复

直接修复是最简单的一种 DNA 损伤修复方式。直接作用于受损的 DNA，将之恢复为原来的结构。

表 4-1　常见 DNA 损伤的修复系统

修复类型	损伤机制	参与修复的酶或蛋白
光复活作用	嘧啶二聚体	光复活酶
碱基切除修复	受损的碱基	DNA 糖基化酶、无嘌呤/无嘧啶核酸内切酶
核苷酸切除修复	嘧啶二聚体、DNA 螺旋结构的改变	大肠杆菌中 UvrA、UvrB、UvrC 和 UvrD，人 XP 系列蛋白 XPA、XPB、XPC……XPG 等
错配修复	复制和重组中的碱基配对错误	大肠杆菌中的 MutH、MutL、MutS，人的 hMLH1、hMSH2、hMSH3、hMSH6 等
重组修复	双链断裂	RecA 蛋白、Ku 蛋白、DNA-PKcs、XRCC4
损伤跨越	大范围的损伤或复制中来不及修复的损伤	RecA 蛋白、LexA 蛋白、其他类型的 DNA 聚合酶

1. 嘧啶二聚体的直接修复　又称为光复活作用，生物体内存在一种光复活酶（photoreactivating enzyme），能够直接识别和结合于 DNA 链上的嘧啶二聚体部位，在波长300～500 nm 的可见光激发下，将其解聚为原来的单体核苷酸从而修复（图 4-3）。光复活酶最初在低等生物中发现，高等生物虽然存在该酶，但是光复活并不是高等生物修复嘧啶二聚体的主要方式。

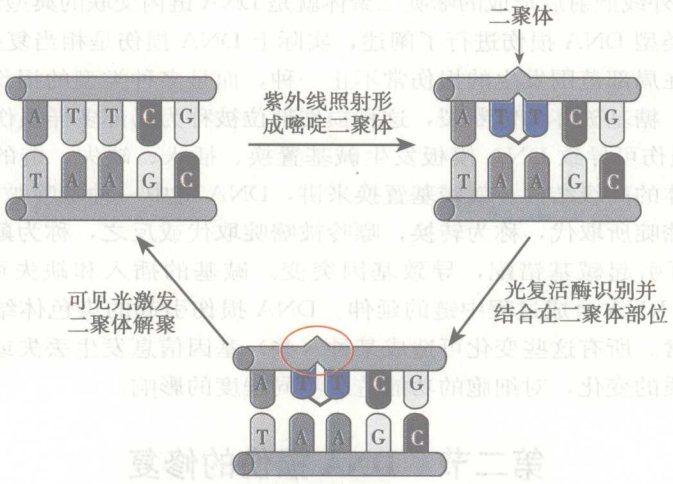


图 4-3　胸腺嘧啶二聚体的光复活修复

2. 烷基化碱基的直接修复　催化此类直接修复的酶是一类特异的烷基转移酶，它可以将烷基从核苷酸转移到自身肽链上，修复 DNA 的同时自身发生不可逆转的失活。如人类 O^6-甲基鸟嘌呤-DNA 甲基转移酶能够将 O^6 位的甲基转移到酶自身的半胱氨酸残基上，使甲基化的鸟嘌呤恢复正常结构（图 4-4）。

3. 无嘌呤位点的直接修复　DNA 链上的嘌呤碱基受到致损因素作用时，可能被糖基化酶水解而脱落，生成无嘌呤位点。DNA 嘌呤插入酶能催化游离嘌呤碱基或脱氧核苷与 DNA 缺嘌呤部位重新生成糖苷共价键，导致嘌呤碱基的直接插入。这种作用具有很强的专一性。

4. 单链断裂的直接修复　DNA 连接酶能催化 DNA 双螺旋结构中一条链上缺口处的 5′-磷酸基团与相邻片段的 3′-羟基之间形成磷酸二酯键，从而直接参与部分 DNA 单链断裂的修复，如电离辐射所造成的切口。

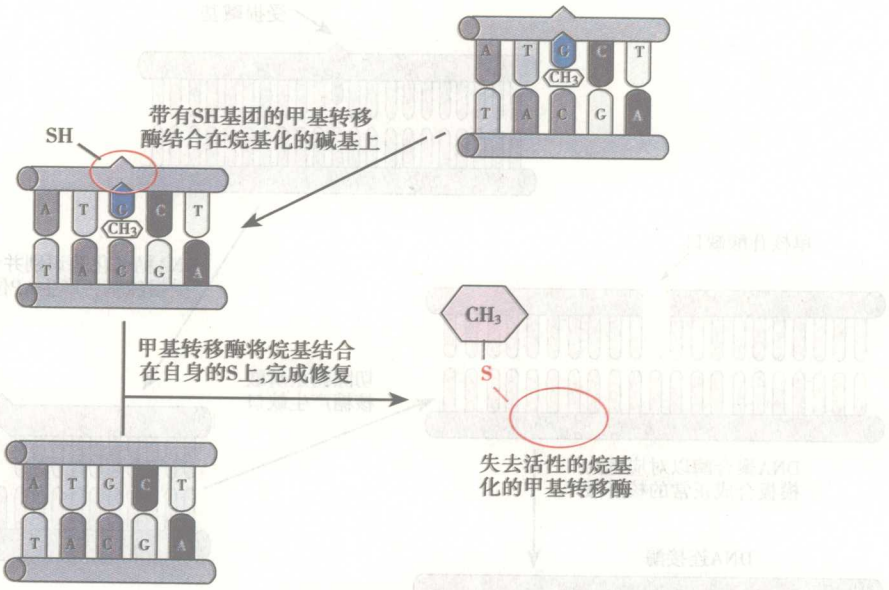


图 4-4　烷基化碱基的直接修复

二、切除修复是最普遍的 DNA 损伤修复方式

切除修复是生物界最普遍的一种修复方式，通过修复系统将不正常的碱基或核苷酸除去并替换掉。依据识别损伤机制的不同，可分为碱基切除修复和核苷酸切除修复两种方式。

1. 碱基切除修复　依赖于生物体内存在的一类特异的 DNA 糖基化酶。整个修复过程包括：①识别水解：DNA 糖基化酶特异性识别 DNA 中已受损的碱基并将其水解去除，产生了一个无碱基位点；②切除：无碱基位点核酸内切酶在此位点 5′端将 DNA 链的磷酸二酯键切开，去除剩余的磷酸糖基部分；③合成：DNA 聚合酶在缺口处以另一条链为模板修补合成互补序列；④连接：由 DNA 连接酶将切口重新连接，使 DNA 恢复正常结构（图 4-5）。

抑癌基因 P53 在哺乳动物细胞中参与调控碱基切除，其直接证据是 DNA 烷化剂诱导的 DNA 损伤在表达野生型 P53 的细胞可被有效修复，而在 P53 缺失的细胞修复速度明显减慢。

2. 核苷酸切除修复　与碱基切除修复不同，核苷酸切除修复系统并不识别具体的损伤，而是识别损伤对 DNA 双螺旋结构造成的扭曲。其修复过程与碱基切除修复相似，首先由一个酶系统识别 DNA 损伤部位，然后在损伤两侧切开 DNA 链，去除两个切口之间的一段受损的寡核苷酸，缺损区在 DNA 聚合酶作用下以另一条链为模板，合成一段新的 DNA 而被填补，最后由连接酶连接。

切除修复是 DNA 修复的一种普遍形式，它并不局限于某种特殊原因造成的损伤，而能一般性地识别和纠正 DNA 链及 DNA 螺旋结构的变化，修复系统能够使用相同的机制和一套修复蛋白去修复一系列性质各异的损伤。

遗传性着色性干皮病（xeroderma pigmentosum，XP）的患者对光线极度敏感而导致皮肤的损伤，包括皮肤癌。目前已知与其基因的突变有关，这些基因以其缩写命名为 XP

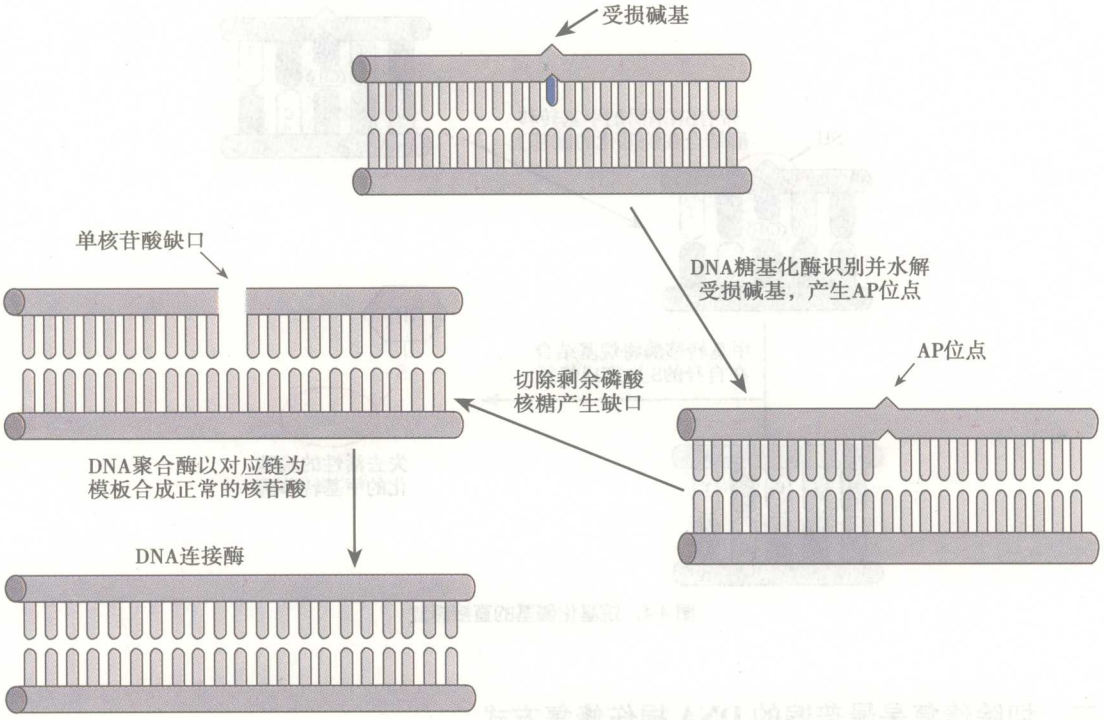


图 4-5　单个碱基的切除修复

基因，表达产生的 XP 系列蛋白有 XPA、XPB、XPC……XPG 等，均参与核苷酸切除修复途径。人的核苷酸切除修复需要 30 多种蛋白的参与。其修复过程如下：①首先由损伤识别蛋白 XPC、XPA 等，再加上复制蛋白 SSB，结合在受损伤 DNA 部位。②XPB、XPD 发挥解旋酶活性，与上述物质共同作用在受损周围形成一个凸起；③XPG 与 XPF 发生构象改变，分别在凸起 3′侧和 5′侧发挥核酸内切酶活性，在增殖细胞核抗原（PCNA）帮助下，切除并释放受损的寡核苷酸；④遗留的缺损区由聚合酶 δ 或 ε 一起进行修补合成，最后连接酶完成连接。此外，还有柯凯氏综合征（Cockayne syndrome）、人毛发二硫键营养不良症（trichothiodystrophy）等。

核苷酸切除修复不仅能够修复整个基因组中的损伤，而且能修复那些正在转录的基因模板链上的损伤，后者又称为转录偶联修复（transcription-coupled repair），更具积极意义。在此修复中，所不同的是由 RNA 聚合酶承担起识别损伤的重任。

3. 碱基错配修复　错配是指非 Watson-Crick 碱基对，错配修复可看作是碱基切除修复的一种特殊形式，它纠正复制和重组中的碱基配对错误及因碱基损伤所致的碱基配对错误、碱基插入、缺失等损伤，是维持 DNA 结构完整稳定的重要方式。从低等生物到高等生物都具有这一修复系统。

大肠杆菌参与 DNA 复制产生的错配修复的蛋白包括 Mut（mutase）H、MutL、MutS、DNA 解旋酶、单链结合蛋白、核酸外切酶Ⅰ、DNA 聚合酶Ⅲ及 DNA 连接酶等 10 余种，修复过程十分复杂。修复过程中错配修复系统面临的主要问题是如何区分母链和子链。在细菌 DNA 中甲基化修饰是一个重要标志，母链是高度甲基化的，其 A 具有甲基修饰，而新合成子链中 A 的甲基化尚未进行，这就提示修复应在此链上进行。首先由 MutS 复合物蛋白识别错配碱基，随后由 MutL、MutH 等蛋白协同相应核酸外切酶将包含错配点在内的一小段 DNA 水解、切除，经修补、连接后，恢复 DNA 正确的碱基配对。

以细菌错配修复系统的研究为基础，真核细胞的错配修复机制的研究也有很大进展。现已发现多种与大肠杆菌 MutS、MutL 高度同源的参与错配修复的蛋白，如与大肠杆菌 MutS 高度同源的 hMSH2（human muscle segment homeobox 2，hMSH2）、hMSH6、hMSH3 等。hMSH2 和 hMSH6 复合物可识别包括碱基错配、插入、缺失等损伤，而由 hMSH2 和 hMSH3 形成的复合物则主要识别碱基插入、缺失等损伤。真核细胞并不像原核生物那样以甲基化来区分母链和子链，可能是依赖修复酶与复制复合体之间的联合作用来识别。

三、DNA 严重损伤时需要进行重组修复

双链 DNA 分子中一条链的断裂，如像碱基和核苷酸切除修复过程中以及某些氧化损伤产生的单链断裂，可被模板依赖的 DNA 合成系统修复，并不会给细胞带来严重的后果。但 DNA 分子的双链断裂是一种极为严重的损伤，与其他修复系统不同的是，双链断裂修复没有互补链来提供修复断裂的遗传信息，需要另外一种更为复杂的机制来指导修复即重组修复。重组修复是指依靠重组酶系将另一段未受损伤的 DNA 移到损伤部位，提供正确的模板，进行修复的过程。依据机制的不同，重组修复可分为同源重组和非同源末端连接重组。

1. 同源重组修复　所谓同源重组（homologous recombination）修复指的是参加重组的两段双链 DNA 在相当长的范围内顺序相同（≥200 bp），这样能保证重组后生成的新区顺序正确。大肠杆菌和酵母同源重组的分子机制比较清楚，起关键作用的是 RecA 蛋白，也称作重组酶，它是一个由 352 个氨基酸组成的蛋白。多个 RecA 单体在 DNA 上聚集可形成右手螺旋的核蛋白细丝，细丝中具有深的螺旋凹槽，可以识别和容纳 DNA 链。在 ATP 存在的情况下，RecA 可与损伤的 DNA 单链区结合，使 DNA 伸展，同时 RecA 可识别一段与受损 DNA 序列相同的姐妹链，并使之并列，进而在姐妹链的母链上切口，两条链间发生重组交换。结构正常的母链 3′端侵入待修复的缺口处，以此为模板重建损伤链。因交换形成的缺损可以另一链为模板合成修复。最后在其他酶的作用下切开交叉连接处，连接新合成的链（图 4-6）。同源重组生成的新片段具有很高的忠实性。

2. 非同源末端连接的重组修复　顾名思义，即两个分子的末端不需要同源性就能连接起来，因此非同源末端连接重组（non-homologous end joining，NHEJ）修复的 DNA 链同源性不高，修复的 DNA 序列存在一定的错误，但对于具有巨大基因组的哺乳动物细胞来说，发生错误的位置可能并不在必需基因上，可以维持细胞的存活，是哺乳动物细胞 DNA 双链断裂的修复方式。非同源重组中起关键作用的蛋白分子是 DNA 依赖的蛋白激酶（DNA-dependent protein kinase，DNA-PK）。它是一种核内的丝氨酸/苏氨酸蛋白激酶，由一个分子量大约为 465 kDa 的催化亚基（DNA-PKcs）和一个能结合 DNA 游离端的杂二聚体蛋白 Ku 组成。DNA-PKcs 的主要作用是介导 DNA-PK 的催化功能，Ku 蛋白可与双链 DNA 的断端连接，促进双链断裂的重接与重组。

另一个参与非同源重组的重要蛋白是 XRCC4（X-ray repair，complementing defective，in Chinese hamster），它能与 DNA 连接酶形成复合物，并增强连接酶的活力，在 DNA 连接酶与组装在 DNA 末端的 DNA-PK 复合物相结合的过程中起中间体作用。NHEJ 既是修复 DNA 损伤的方式，又可以看作是一种生理性重组策略，将原来并未连在一起的分子或片段连接产生新的组合，如 B 淋巴细胞、T 淋巴细胞受体基因、免疫球蛋白基因的构建与重排等。

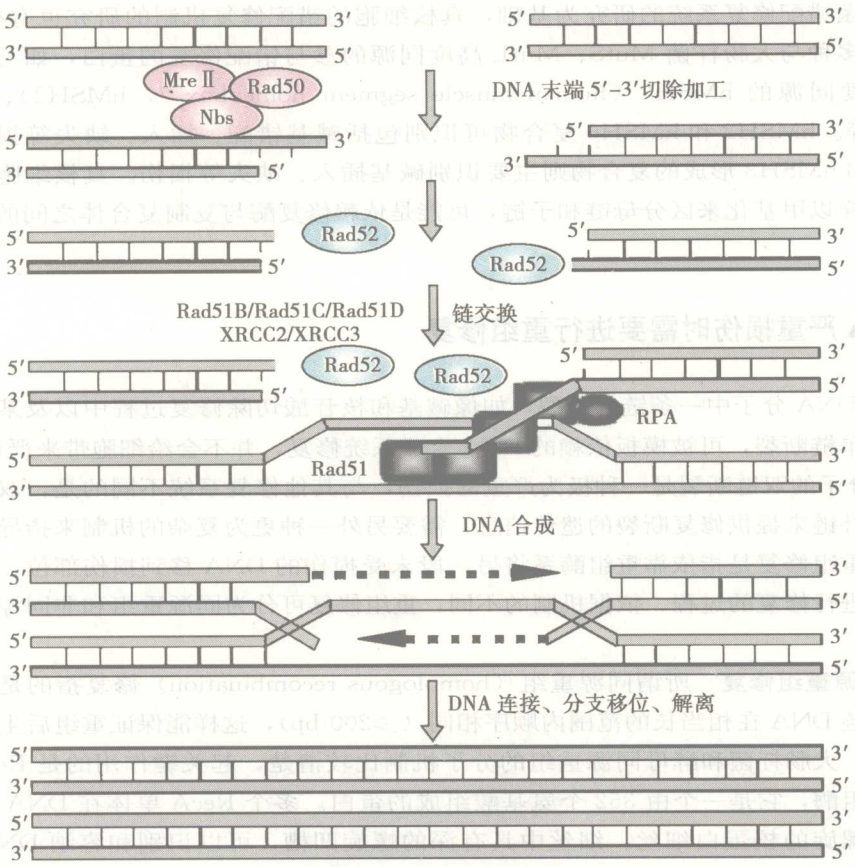


图 4-6　同源重组修复

四、跨越 DNA 损伤的复制后修复

当 DNA 双链发生大范围的损伤或是复制叉已经解开母链，致使修复系统无法通过上述的多种方式进行修复，DNA 损伤部位失去了模板作用而中止复制时，细胞可以诱导一个或多个应急途径，跨过损伤部位先进行复制，再设法修复。研究最多的诱导途径是 SOS 反应。根据损伤部位跨越机制的不同分为重组跨越和损伤合成跨越。

1. 重组跨越损伤部位　当 DNA 链的损伤较大致使损伤链不能作为模板复制时，细胞利用同源重组的方式将 DNA 模板进行重组交换，使复制能够继续下去。大肠杆菌中，当复制进行到损伤部位时，DNA 聚合酶Ⅲ停止移动从模板上脱离下来，然后在损伤下游重新启动复制，从而在子链 DNA 上产生一个缺口。RecA 重组蛋白将另一股健康母链上对应序列重组到子链 DNA 的缺口处填补。通过重组跨越，解决了有损伤的 DNA 分子的复制问题，但其损伤并没有真正地修复，只是转移到新合成的一个子代 DNA 分子上，由细胞内其他修复系统来后继修复或是在不断复制之中被“稀释”掉。

2. 跨越损伤的 DNA 合成　DNA 双链发生大片段、高频率的损伤时，细胞可以紧急启动应急修复系统，诱导产生新的 DNA 聚合酶，替换停留在损伤位点的原来 DNA 聚合酶Ⅲ，在子链上以随机方式插入正确或错误的核苷酸使复制继续，越过损伤部位之后，这些新 DNA 聚合酶完成使命从 DNA 链上脱离，再由原来的 DNA 聚合酶Ⅲ继续复制。因为诱导产生的这些新 DNA 聚合酶的活性低，识别碱基的精确度差，一般无校对功能，所以

这种损伤合成跨越的复制过程的出错率会大大增加。是大肠杆菌 SOS 反应或 SOS 修复的一部分。

在原核细胞中 SOS 修复反应是由 RecA 蛋白和 LexA 阻遏物相互作用引起的，近 30 个“*sos*”基因参与反应。正常情况下 RecA 基因及其他“*sos*”可诱导基因的上游有一段共同的操纵序列（5′-CTG-N10-CAG-3′）被 LexA 阻遏蛋白所阻遏抑制，只有低水平转录和翻译，产生少量蛋白。当 DNA 严重损伤时，RecA 蛋白被激活而具有了促 LexA 水解的酶活性，当 LexA 阻遏蛋白因水解而从“*sos*”基因的操纵序列上解离下来后，一系列原本受抑制的“*sos*”基因得以表达，从而参与 SOS 修复活动。完成修复后，LexA 阻遏蛋白被重新合成，相关“*sos*”可诱导基因又被关闭（图 4-7）。SOS 反应诱导的产物可参与重组修复、切除修复、错配修复等各种途径的修复过程。这种修复机制因海空紧急呼救信号“SOS”而得名。

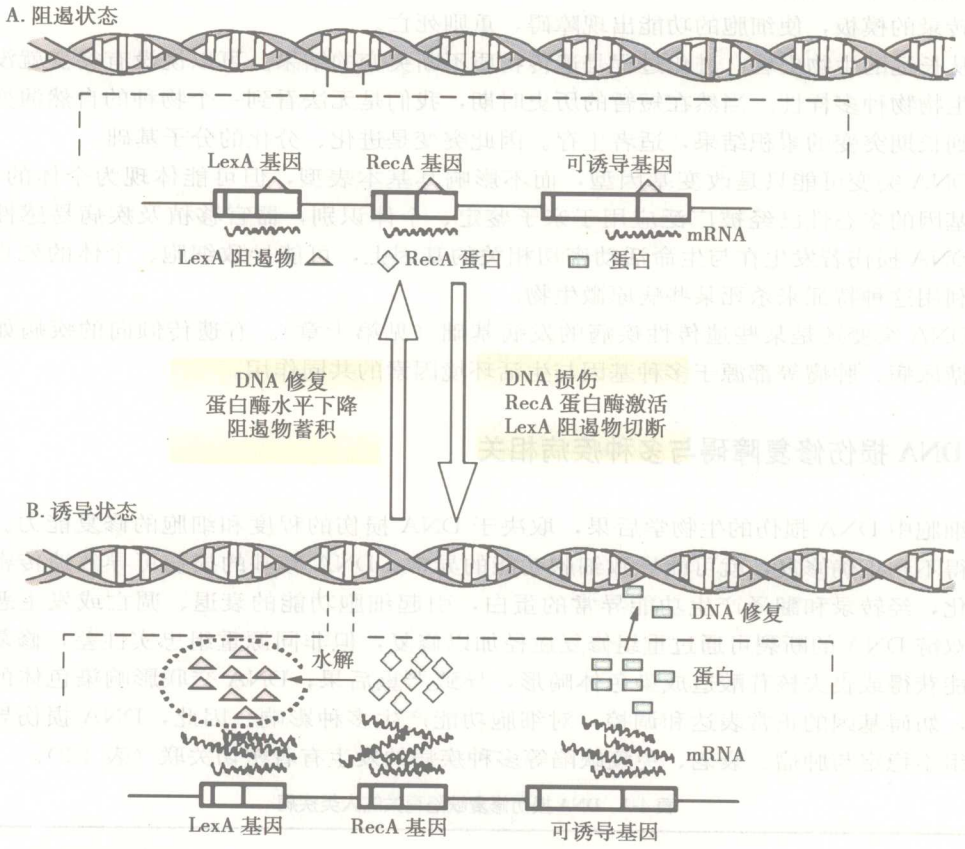


图 4-7　SOS 修复中 LexA-RecA 操纵子的作用机理

此外，对于受损的 DNA 分子，细胞除了启动上述诸多的修复系统修复损伤之外，还可以通过其他途径将损伤的后果降至最低，如通过 DNA 损伤应激反应活化的细胞周期检查点机制，延迟或阻断细胞周期进程，为损伤修复提供时间，使细胞能安全进入新一轮细胞周期；诱导修复基因的转录，加强损伤的修复；激活细胞凋亡机制，诱导严重损伤的细胞凋亡等维持生物体基因组的稳定。

第三节　DNA 损伤和修复的意义

遗传物质稳定性的世代相传是维持物种相对稳定的主要因素，但是如果遗传物质是绝对一成不变的话，自然界就会失去进化的基础，也就没有新的物种出现。因此生命和生物多样性依赖于 DNA 损伤（突变）与损伤修复机制的良好的动态平衡。

一、DNA 损伤具有双重效应

一般认为 DNA 损伤都是有害的，但就结果而言，既有消极的一面，也有积极的一面。DNA 损伤之后通常有两个生物学后果：一是给 DNA 带来永久性的改变即突变，可能改变基因的编码序列或者基因的调控序列；二是 DNA 的某些化学改变使得 DNA 不能用作复制和转录的模板，使细胞的功能出现障碍，重则死亡。

从长远的生物史看，进化过程是遗传物质不断突变的结果。可以说没有突变就没有如今的生物物种多样性。当然在短暂的历史时期，我们是无法看到一个物种的自然演变，只能见到长期突变的累积结果，适者生存。因此突变是进化、分化的分子基础。

DNA 突变可能只是改变基因型，而不影响其基本表型，但可能体现为个体的差异。例如基因的多态性已经被广泛应用于亲子鉴定、个体识别，器官移植及疾病易感性分析等。DNA 损伤若发生在与生命活动密切相关的基因上，可能导致细胞、个体的死亡。人类常利用这种特征来杀死某些病原微生物。

DNA 突变还是某些遗传性疾病的发病基础（见第十章）。有遗传倾向的疾病如高血压、糖尿病、肿瘤等都源于多种基因与生活环境因素的共同作用。

二、DNA 损伤修复障碍与多种疾病相关

细胞中 DNA 损伤的生物学后果，取决于 DNA 损伤的程度和细胞的修复能力。如果损伤得不到正确修复，就可能导致细胞功能的异常。DNA 碱基的损伤可导致遗传密码子的变化，经转录和翻译产生功能异常的蛋白，引起细胞功能的衰退、凋亡或发生恶性转化；双链 DNA 的断裂可通过重组修复途径加以修复，但非同源重组忠实性差，修复过程中可能获得或丧失核苷酸造成染色体畸形，导致严重后果；DNA 交联影响染色体的高级结构，妨碍基因的正常表达和调控，对细胞功能产生多种影响。因此，DNA 损伤导致的基因组不稳定与肿瘤、衰老、免疫缺陷等多种疾病的发生有着密切关联（表 4-2）。

表 4-2　DNA 损伤修复缺陷有关的人类疾病

疾病	易患肿瘤或疾病	有缺陷的修复系统
着色性干皮病	皮肤癌、黑色素瘤	核苷酸切除修复
遗传性非息肉病性结肠癌	结肠癌、卵巢癌	错配修复 转录偶联修复
柯凯氏综合征	视网膜萎缩、侏儒、耳聋、早衰、对 UV 敏感	核苷酸切除修复、转录偶联修复
毛发硫营养不良症	毛发易断、生长迟缓	核苷酸切除修复
布卢姆综合征	白血病、淋巴瘤	非同源末端连接重组
范可尼贫血	再生障碍性贫血、白血病、生长迟缓	重组跨越
遗传性乳腺癌	乳腺癌、卵巢癌	同源重组修复

1. DNA 损伤修复缺陷与肿瘤　先天性 DNA 损伤修复缺陷疾病患者容易发生各种恶性肿瘤，肿瘤发生是 DNA 损伤对机体的远后效应之一。众多研究表明，DNA 损伤→DNA 修复异常→基因突变→肿瘤发生是贯穿肿瘤发生过程中的关键环节。DNA 损伤可导致原癌基因的激活，也可使抑癌基因失活，癌基因、抑癌基因的失衡是细胞恶变的重要机制。参与 DNA 修复的多种基因具有抑癌基因的功能，目前已发现这些基因在多种肿瘤中发生突变而失活。1993 年，研究发现人类遗传性非息肉病性结肠癌（hereditary non-polyposis colorectal cancer，HNPCC）细胞存在错配修复、转录偶联修复缺陷，造成细胞基因组的不稳定性，进而引起调控细胞生长基因的突变，发生细胞恶变。

BRCA 基因（breast cancer gene）参与 DNA 损伤修复的启动，细胞周期的调控，*BRCA* 的失活可增加细胞辐射敏感性，并导致双链断裂修复能力的下降。现已发现 *BRCA*1 基因在 70%的家族遗传性乳腺癌和卵巢癌病例中发生突变而失活。

值得注意的是，DNA 修复功能缺陷虽可引起肿瘤的发生，但已癌变的细胞本身 DNA 修复功能往往并不低下，相反却显著地升高，使得癌细胞能够充分修复化疗药物引起的 DNA 损伤，这也是大多数抗癌药物不能奏效的原因，所以 DNA 修复的研究可为肿瘤联合化疗提供方案。

2. DNA 损伤修复缺陷与人类遗传病　着色性干皮病患者的皮肤对日光敏感，照射后出现红斑、水肿，继而出现色素沉着、干燥、角化过度，结果可导致黑色素瘤、基底细胞癌、鳞状上皮癌及棘状上皮瘤等肿瘤的发生。具有不同临床表现的该病患者存在明显的遗传异质性，表现为不同程度的核酸内切酶缺乏引起的切除修复功能缺陷，所以患者的肺、胃肠道等器官在受到环境等因素刺激下，将有较高的肿瘤发生率。在对 *XP* 的研究中，还发现一些患者虽具有明显的临床症状，但在 UV 辐射后的核苷酸切除修复中却没有明显的缺陷表型，将其定名为“*XP* 变种”（XP variant，XPV）。这类患者的细胞在培养中表现对 UV 辐射的轻微增高的敏感性，变种的切除修复功能正常，但复制后修复的功能有缺陷。最新的研究发现某些 *XP* 变种的分子病理学机制是由于它对 DNA 碱基损伤耐受的缺陷所致，而不是修复方面的缺陷。

此外 DNA 损伤修复机能的缺陷还可以导致毛发硫营养不良症（trichothiodystrophy，TTD）、柯凯氏综合征（Cockayne syndrome，CS）、范可尼贫血（Fanconi anemia）等遗传病。

3. DNA 损伤修复与衰老　从 DNA 修复功能的比较研究中发现寿命长的动物，如象、牛等的修复能力较强；寿命短的动物如小鼠、仓鼠等的修复能力较弱。人的 DNA 修复能力也很强，但到一定年龄后逐渐减弱，突变细胞数、染色体畸变率同时也相应增加。如人类常染色体隐性遗传的早老症和韦尔纳氏综合征患者一般早年死于心血管疾病或恶性肿瘤，患者的体细胞极易衰老。

4. DNA 损伤修复缺陷与免疫　DNA 修复功能先天缺陷的病人的免疫系统亦常有缺陷，主要是 T 淋巴细胞功能的缺陷。随着年龄的增长细胞中的 DNA 修复功能逐渐衰退，如果同时发生免疫监视机能的障碍，便不能及时清除癌化的突变细胞，从而导致发生肿瘤。因此，衰老、DNA 修复、免疫和肿瘤是紧密关联的。

5. DNA 损伤应答缺陷与人类遗传病　共济失调-毛细血管扩张症（Ataxia telangiectasia，AT），是一种常染色体隐性遗传病，主要影响机体的神经系统、免疫系统和皮肤。AT 患者的细胞对射线及拟辐射的化学因子，如博来霉素等敏感，具有极高的染色体自发畸变率，以及对辐射所致 DNA 损伤的修复缺陷。患者的肿瘤发病率相当高。AT 的发生与在 DNA 损伤的信号转导网络中起关键作用的 ATM 分子的突变有关。

（吕建新）

第五章　基因表达

关键知识点

- 基因表达是指基因所贮存的遗传信息通过 RNA 转录和蛋白质生物合成产生具有生物功能的 RNA 或蛋白质的过程。
- 基因转录合成 RNA 的过程是以 DNA 为模板的 RNA 聚合酶所催化的核苷酸聚合反应。RNA 链的合成是从 5′末端开始，向 3′端延伸；对 DNA 模板链的阅读沿 3′端向 5′端方向进行。
- 原核生物只有 1 种 RNA 聚合酶，全酶由 5 个亚基构成的核心酶（$\alpha_2\beta'\beta\omega$）和 1 个 σ 因子组成，负责所有种类 RNA 的合成；真核生物则有 RNA 聚合酶Ⅰ、RNA 聚合酶Ⅱ和 RNA 聚合酶Ⅲ等 3 种 RNA 聚合酶，分别负责转录合成不同种类的 RNA，其中 RNA 聚合酶Ⅱ负责 mRNA 的合成。
- RNA 均在基因的启动子中的特定起始位点开始合成。在原核生物，RNA 聚合酶中的 σ 因子负责识别起始位点；真核生物的基因转录起始需要除 RNA 聚合酶外的其他转录因子共同形成起始复合物。与原核生物相比，真核生物 RNA 合成后需要更为复杂的加工修饰，如 mRNA 前体的 5′端加帽、3′端加尾、剪接去除内含子序列以及 RNA 编辑等。
- 细胞内蛋白质的合成以 mRNA 为模板进行，mRNA 从 5′至 3′方向的密码子排列决定着合成肽链的从 N 端至 C 端的氨基酸排列顺序。由多种蛋白质和 rRNA 构成的核糖体是肽链合成的场所。真核生物和原核生物蛋白质合成机制相似，但过程更为复杂。

关键词： 基因表达；转录；RNA 聚合酶；翻译；遗传密码；氨基酰 tRNA 合成酶；核糖体循环

细胞的生命活动由蛋白质或 RNA 所完成。基因组携带的固有遗传信息指导下生成各种具有生物学功能的 RNA 或蛋白质的过程（图 5-1）称为基因表达（gene expression），是基因所携带的遗传信息呈现出表型的过程。以 DNA 为模板合成 RNA 的过程称为转录（transcription）；以 mRNA 为模板合成蛋白质的过程称为翻译（translation）。真核生物的 RNA 转录在细胞核内完成，而翻译合成蛋白质则在细胞质中进行，故合成的 RNA 需要转运至胞质。原核生物没有细胞核，遗传信息的转录与翻译是同步进行的。

图 5-1　DNA 转录与翻译

第一节　RNA 的转录合成

基因转录是在 RNA 聚合酶催化下，以 ATP、GTP、CTP、UTP 为基本原料单位，以 DNA 为模板依碱基配对原则（G/C、C/G、T/A、A/U）合成 RNA 的过程。基因转录产物包括 mRNA、tRNA、rRNA 及其他小分子 RNA。各类 RNA 合成的生物化学过程基本相同，不同之处主要表现在参与合成的 RNA 聚合酶的种类、转录调节方式、转录后加工过程等方面。

一、基因组 DNA 是 RNA 合成的模板

DNA 链是有极性的。RNA 聚合酶以不对称的方式与启动子结合，使得转录只能沿着一个方向进行，一个基因的两条互补链中也只能有一条被转录成 RNA。这种现象称为不对称转录（asymmetric transcription）。双链 DNA 中作为 RNA 合成模板的一条链称为模板链（template strand）；相互补的另一条链是非模板链。新合成的 RNA 序列与模板链 DNA 序列反向互补、而与非模板链相同（U 替代 T），故非模板链亦被称为编码链（coding strand）。不同基因的模板链与非模板链并非固定在 DNA 双螺旋中某一条链上，即一个 DNA 分子中的不同基因的模板链可以位于该分子的不同单链上（图 5-2）。因此位于 DNA 不同单链上的基因转录方向将不同。

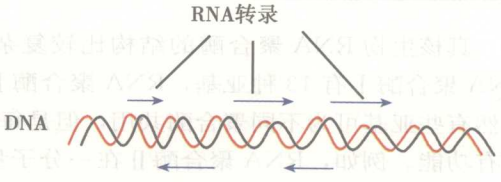


图 5-2 不对称转录示意图

二、DNA 指导的 RNA 聚合酶催化 RNA 合成

催化基因转录的酶是 DNA 指导的 RNA 聚合酶（DNA-dependent RNA polymerase），简称 RNA 聚合酶。在 RNA 链聚合反应中，一个核苷酸分子的 5′-三磷酸的 α-磷酸基团与另一个核苷酸分子的 3′-OH 聚合，释放出 1 分子焦磷酸，形成 3′，5′磷酸二酯键，焦磷酸进一步水解产生 2 分子无机磷酸，产生的能量不断推动反应进行。RNA 链的合成方向是从 5′→3′，即新的核苷酸分子总是加在原有 RNA 链的 3′-OH 末端。

$$\underset{\text{RNA}}{(\text{NMP})_n}+\text{NTP}\xrightarrow{\text{RNA 聚合酶}}\underset{\text{延长的 RNA}}{(\text{NMP})_{n+1}}+\text{PP}_i$$

以 DNA 为模板，转录完成后得到的 RNA 分子称为初级转录物（primary transcript），初级转录物还需经过加工修饰后方可成为有功能的成熟 RNA。

（一）原核生物只有一种 RNA 聚合酶

在原核生物中迄今只发现了一种 RNA 聚合酶，负责所有种类 RNA（mRNA、tRNA 和 rRNA 等）的合成。目前结构与功能研究得比较清楚的是大肠杆菌的 RNA 聚合酶。大肠杆菌 RNA 聚合酶的分子质量接近 500 kDa，是一个由 5 种亚基组成的复合酶。这 5 种亚基分别是 α、β、β′、ω 和 σ，其中 2 分子 α 亚基和各 1 分子 β、β′、ω 构成核心酶（core enzyme），核心酶（$\alpha_2\beta'\beta\omega$）和 σ 亚基（σ subunit）再构成全酶（holoenzyme）。大肠杆菌的 RNA 聚合酶的 σ 亚基负责识别启动子，并启动转录，而核心酶负责 RNA 链的延伸。核心酶中的 β 和 β′构成酶的催化活性中心，α 亚基参与转录速度的调控，而 ω 亚基的功能尚不清楚。

（二）真核生物有 3 种不同的 RNA 聚合酶

真核生物中至少已经发现了 3 种 RNA 聚合酶，分别称为 RNA 聚合酶Ⅰ、RNA 聚合酶Ⅱ和 RNA 聚合酶Ⅲ。由于这些 RNA 聚合酶所识别的启动子序列不同（见第一章），因而将分别负责不同种类 RNA 的转录合成（表 5-1）。真核生物不同的 RNA 聚合酶之间有较高的结构类似性，且与原核生物 RNA 聚合酶亦有一定相似性，提示 RNA 聚合酶在生物进化过程中可能源于同一祖先且高度保守。

表 5-1　真核生物 RNA 聚合酶的种类和性质

	RNA 聚合酶Ⅰ	RNA 聚合酶Ⅱ	RNA 聚合酶Ⅲ
细胞内定位	核仁	核质	核质
转录产物	45S rRNA	mRNA、miRNA	tRNA、5S rRNA、snRNA

真核生物 RNA 聚合酶的结构比较复杂，都是含有多种不同亚基的复合物。酵母的 RNA 聚合酶Ⅰ有 13 种亚基，RNA 聚合酶Ⅱ有 10 种亚基，RNA 聚合酶Ⅲ有 14 种亚基。虽然有些亚基可为不同聚合酶共用，但是每种 RNA 聚合酶都有 4～7 个特有亚基，决定其特有功能。例如，RNA 聚合酶Ⅱ在一分子量为 220 kDa 的亚基上有一特殊的羧基末端结构域（carboxyl terminal domain，CTD），为该亚基所独有。该结构域上的丝氨酸和苏氨酸残基的磷酸化或去磷酸化直接影响 RNA 聚合酶Ⅱ的活性。

RNA 聚合酶与 DNA 聚合酶相比较在催化作用上有两个重要的不同：①RNA 聚合酶不需要引物就能直接启动核苷酸的聚合反应，而大部分 DNA 聚合酶需要 RNA 引物存在方可启动合成；②RNA 聚合酶不具有 3′→5′核酸外切酶活性，故不具有校读作用。因此转录的忠实性要比 DNA 复制过程低，不过由于 RNA 的错误合成并不影响子代细胞的性状，因此少量的错误转录产物对生物不会产生重大影响。

三、转录起始复合体在启动子部位形成

为便于理解，可以人为地将 RNA 合成过程分为起始、延长和终止 3 个阶段。在起始阶段，RNA 聚合酶首先需要识别双链 DNA 模板中的转录起始部位，即启动子区域，与之结合形成转录起始复合体并启动 RNA 链的延伸。

（一）原核生物 RNA 聚合酶的 σ 亚基负责启动子识别

原核生物 RNA 合成起始阶段需要 RNA 聚合酶全酶（图 5-3）。RNA 聚合酶全酶与 DNA 以较低的亲和力疏松结合着，可以沿着 DNA 双链滑动，当 σ 亚基识别到启动子序列时，RNA 聚合酶在此部位紧密结合在 DNA 模板上，形成转录起始复合体。此时 DNA 在结构上仍处于双链状态，故该复合体被称为闭合转录复合体（closed transcription complex）。启动子中的 Pribnow 盒序列（见第一章）中富含 AT 碱基，因而氢键较易断裂。RNA 聚合酶打开启动子部位的 DNA 双螺旋两条链之间的氢键使 DNA 发生约 17 bp 长的局部解链，形成开放转录复合体（open transcription complex），暴露出 DNA 模板链。在转录起始部位形成新合成的 RNA 第一个磷酸二酯键，RNA 链的第一个核苷酸均为嘌呤核苷酸 A 或 G。

（二）真核生物 RNA 聚合酶启动转录需要其他蛋白质因子

真核生物的转录起始远比原核生物复杂，除了 RNA 聚合酶以外，还需要有多种蛋白因子参与方可形成转录起始复合物。这类因子属于一大类被称为转录因子（transcription factor，TF）的蛋白质。除了起始阶段以外，在 RNA 链的延长阶段也同样需要转录因子的参与。RNA 聚合酶Ⅰ、Ⅱ、Ⅲ分别催化不同种类 RNA 的合成，需要结合在不同的启动子部位（见第一章），因而各自有相对应的三类转录因子存在，分别称为 TFⅠ、TFⅡ和 TFⅢ。这里以 RNA 聚合酶Ⅱ为例，介绍真核生物转录起始复合物的形成和转录的起始。

RNA 聚合酶Ⅱ负责 mRNA 前体 hnRNA 的合成。参与 RNA 聚合酶Ⅱ起始的转录因子有 TATA 盒结合蛋白（TATA box-binding protein，TBP）、TFⅡA、TFⅡB 等多种

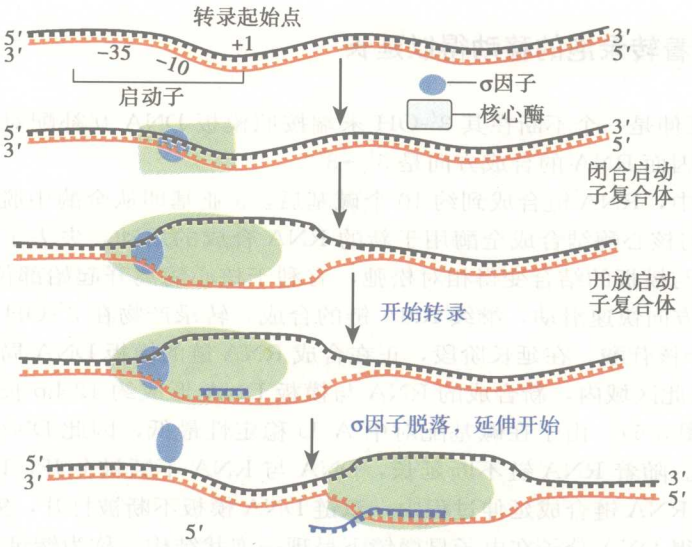


图 5-3　大肠杆菌 RNA 聚合酶转录起始过程

（表 5-2）。

表 5-2　RNA 聚合酶Ⅱ相关转录因子

转录因子	功　能
TBP	识别并结合 TATA 盒
TFⅡA	稳定 TBP 和 TFⅡB 与启动子区域的结合
TFⅡB	结合 TBP、结合聚合酶Ⅱ-TFⅡF 复合体
TFⅡE	有解旋酶和 ATP 酶活性、与 TFⅡH 结合
TFⅡF	与聚合酶Ⅱ形成复合体，结合 TFⅡB，维持聚合酶Ⅱ的特异性
TFⅡH	具有打开 DNA 的解旋酶活性，使聚合酶Ⅱ的 CTD 发生磷酸化

真核生物的转录起始复合体形成过程（图 5-4）可以归纳为两步：

1. 形成闭合转录复合体　首先 TBP 识别 DNA 模板的 TATA 盒并结合于启动子区域；TFⅡB 结合于 TBP 及 DNA；TFⅡF 与 RNA 聚合酶Ⅱ的复合物到达起始部位形成一前起始复合物；具有结合 TFⅡH 能力的 TFⅡE 结合到前起始复合物，募集 TFⅡH 以完成解链准备。

2. 形成开放转录复合体　在 ATP 存在下，具有解旋酶活性的 TFⅡH 在转录起始位点局部解开 DNA 双链，形成开放复合体。TFⅡH 还具有激酶活性，可以使 RNA 聚合酶Ⅱ羧基末端结构域（CTD）发生多位点磷酸化。CTD 磷酸化导致开放转录复合体构象改变，启动转录。在各种 NTP 存在条件下 RNA 聚合酶Ⅱ开始转录模板链，合成新的 PNA 链。

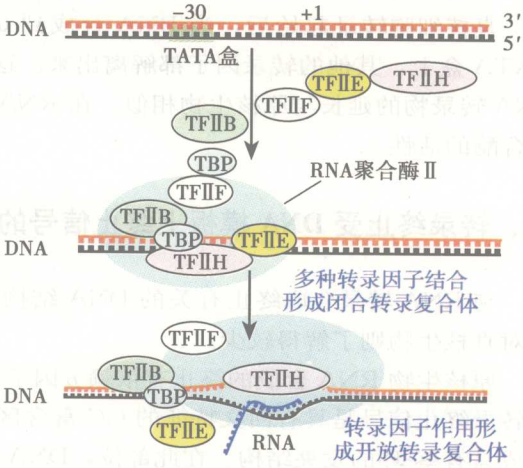


图 5-4　真核生物 RNA 聚合酶Ⅱ参与的起始复合物形成

四、RNA 链随着转录泡的移动得以延长

RNA 链的延伸是一个不断在其 3′-OH 末端按照模板 DNA 互补配对原则加入新的单核苷酸的过程，因而 RNA 的合成方向是 5′→3′。

在大肠杆菌中，RNA 链合成到约 10 个碱基后，σ 亚基即从全酶中脱落。脱离下来的 σ 亚基可以再次与核心酶结合成全酶用于新的 RNA 合成的启动。失去了 σ 亚基的核心酶构象发生改变，与模板的结合变得相对松弛，有利于核心酶离开起始部位，在 DNA 模板上沿着 3′→5′的方向快速滑动，继续 RNA 链的合成。转录产物在 3′-OH 末端延长速度约为每秒 30～85 个核苷酸。在延长阶段，正在合成 RNA 链的模板 DNA 局部双链解开约 17 bp 长的区域。在此区域内，新合成的 RNA 与模板 DNA 形成约 12 bp 长的局部杂化螺旋（hybrid helix）（图 5-5）。由于在碱基配对中 A/U 稳定性最低，因此 DNA-RNA 的杂交链的氢键较易分开。随着 RNA 链不断延长，DNA 与 RNA 不断被分开，DNA 双螺旋结构又得以恢复。在 RNA 链合成延伸过程中，双链 DNA 模板不断被打开，完成 RNA 合成后又恢复的现象使得 DNA 分子在电子显微镜下呈现一泡状结构，称为转录泡（transcription bubble）。转录泡由局部解链的 DNA 模板、RNA 聚合酶及新生的 RNA 构成。

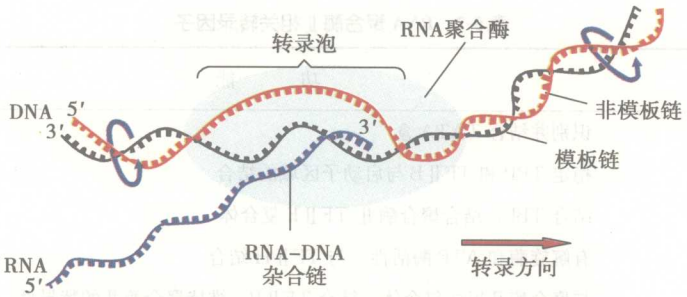


图 5-5　大肠杆菌 RNA 聚合酶转录延长过程的结构变化

在电子显微镜下观察原核生物的转录，可在正在合成的 RNA 链上见到一些点状物，这些点状物被证明是核糖体。可见在原核生物中，转录尚未完成，已在 RNA 链上开始进行蛋白质的翻译了。这一现象称为转录与翻译偶联。

真核细胞转录起始后，当 RNA 合成到 60～70 个核苷酸时，除了 TBP 仍然结合在 TATA 盒上，其他的转录因子都解离出来。这时 RNA 聚合酶Ⅱ进入延长阶段。真核生物 RNA 转录物的延长与原核生物相似。在 RNA 转录延长阶段，延长因子大大增强了 RNA 聚合酶的活性。

五、转录终止受 DNA 模板上终止信号的控制

与 RNA 转录合成终止有关的 DNA 结构和终止机制在原核生物中了解得比较深入，而对真核生物则了解得较少。

原核生物 RNA 合成的终止有依赖 ρ 因子和不依赖 ρ 因子的两种方式（图 5-6）。大多数转录终止信号是具有回文特征的 GC 富含区及 AT 富含区，由此合成的 RNA 转录物能形成具有茎-环的发夹结构。在此部位，DNA 和 RNA 杂化链很不稳定，促使 RNA 链从模板上脱落，从而停止转录。在具有上述终止信号的 DNA 模板上合成 RNA 时，RNA 的合

成将在终止信号处自行终止，不需要终止因子（ρ因子）的帮助。在不具备上述终止信号的DNA转录终止区，最后合成的RNA则含有特殊的ρ因子结合序列，依赖ρ因子的结合而终止RNA的合成。

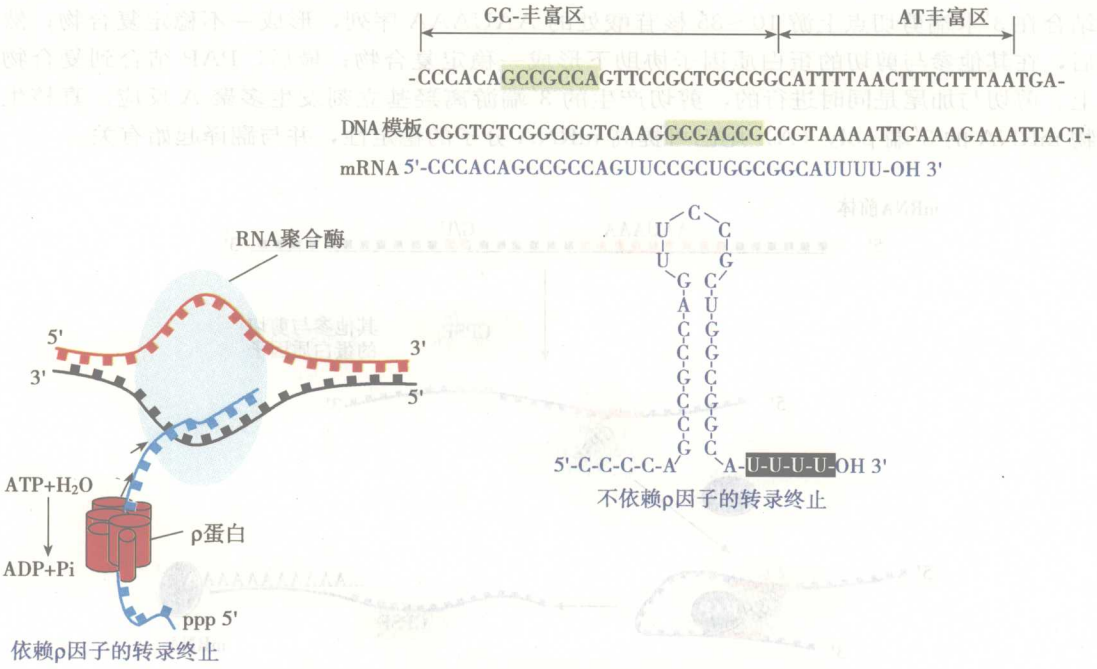


图5-6 原核生物转录的终止

六、初级转录物经转录后加工方可成熟

以DNA为模板合成的RNA分子必须经历一系列转录后的加工和修饰方可成熟为有功能的终产物。原核生物和真核生物的tRNA与rRNA均以前体形式合成初级转录物，这些初级转录物经核酸酶切除一些多余序列，或在某些碱基发生化学修饰，产生成熟的tRNA或rRNA。真核生物的mRNA的前体大小不等，称为不均一核内RNA（heterogeneous nuclear RNA，hnRNA），其转录后加工修饰较为复杂，而原核生物的mRNA合成的同时即被翻译，无需复杂的加工修饰。本部分主要以真核细胞中的mRNA为例，介绍转录后加工修饰基本过程。

（一）mRNA的5′末端帽结构在转录起始后迅速形成

新生mRNA分子长度达到25～30核苷酸时，即在加帽酶催化下，在其5′端加上一分子7甲基鸟苷酸。加帽酶由两个亚基组成，其中一个亚基催化mRNA链5′端的第一个核苷酸（A或G）的γ位磷酸基团水解，形成一个二磷酸核苷残基；该残基接受由加帽酶的另一个亚基所转移的GTP，产生一特殊的5′,5′-三磷酸酯键。在此基础上，在鸟嘌呤甲基转移酶的作用下，由S-腺苷甲硫氨酸提供甲基，G发生N_7甲基化，形成m^7GpppNp-帽结构。5′端帽结构可避免mRNA受到磷酸酶和核酸酶的攻击，维持其结构稳定；可为mRNA前体的剪接提供识别信号；能与帽结合蛋白结合，参与启动蛋白质生物合成。

（二）poly（A）聚合酶负责为mRNA加上3′poly（A）尾

在真核细胞中除了组蛋白mRNA以外，几乎所有的mRNA都有80～250个腺苷酸构

成的 3′末端 poly（A）尾结构。3′末端 poly（A）尾结构的形成是一 3′末端剪切和多聚腺苷酸化同时进行的加工过程（图 5-7）。剪切和多聚 A 作用特异因子（cleavage and polyadenylation specificity factor，CPSF）在细胞中负责 mRNA 前体的 3′末端剪切；poly（A）聚合酶［poly（A）polymerase，PAP］负责在剪切点处添加多聚腺苷酸尾。首先，CPSF 结合在 3′末端剪切点上游 10～35 核苷酸处的 AAUAAA 序列，形成一不稳定复合物；然后，在其他参与剪切的蛋白质因子协助下形成一稳定复合物；最后，PAP 结合到复合物上。剪切与加尾是同时进行的，剪切产生的 3′端游离羟基立刻发生多聚 A 反应。真核生物 mRNA 的 3′端 poly（A）尾巴可提高 mRNA 分子的稳定性，并与翻译起始有关。

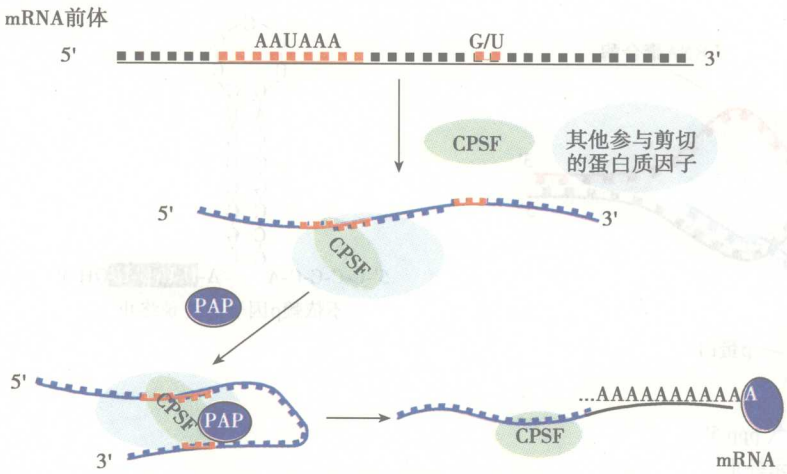


图 5-7　真核生物 mRNA 前体 3′端剪切与加 poly（A）尾过程

（三）mRNA 前体通过剪接切除内含子编码的序列

真核细胞 mRNA 前体（hnRNA）中的内含子序列并不存在于成熟 mRNA 分子中，故切除内含子编码的序列是一重要的转录后加工步骤。内含子剪接部位（splice site）具有特征性结构（图 5-8）。内含子通常在 5′端以 GU 开始，3′端以 AG 结束，并在 3′端上游约 20～50 个核苷酸处存在一分支点。真核生物分支点在 YNCUG 区 C 序列的 A 上（其中 N 代表四种核苷酸中的任何一种，Y 代表 U 或者 C）。内含子剪接过程中会出现一套索结构的中间体，由内含子 5′端的第一个 G 以 2′-5′磷酸二酯键和内含子上的分支点 A 连接而成。剪接过程是两步转酯键反应。首先 5′位点剪断并形成套索状中间体；然后 3′位点剪断和外显子连接同步进行，内含子以套索状结构被切除。

内含子剪接在剪接体（spliceosome）中完成。剪接体由数种蛋白质和 snRNA 形成的 snRNP（见第一章）及其他一些蛋白质共同构成。细胞核内参与内含子剪接的 U1、U2、U4、U5 和 U6 等 5 种 snRNAs 分别与相应蛋白质（共计 50 余种）构成 5 种 snRNPs。这些 snRNPs 通过蛋白质与蛋白质、RNA 与 RNA、RNA 与蛋白质的相互识别，有序地组装到 mRNA 前体的内含子剪接部位上完成剪接。

剪接的基本过程包括：①U1 snRNP 通过 U1 snRNA 识别 5′剪接位点的部分配对碱基而结合于 mRNA 前体；②U2snRNP 结合于 mRNA 前体，U2 snRNA 与分支点处有碱基配对，但是分支点 A 本身与 U2 snRNA 不配对从而形成一突起；③U4/U6/U5 snRNP 通过 U6 snRNA 与 U4 snRNA 5′末端的配对碱基及蛋白质间相互作用形成复合物，再结合于 mRNA 前体而完成剪接体的装配；④snRNAs 和内含子之间、不同的 snRNAs 之间发生重排，催化两个转酯键反应。第二次转酯键反应以后，连接起来的外显子从剪接体中释

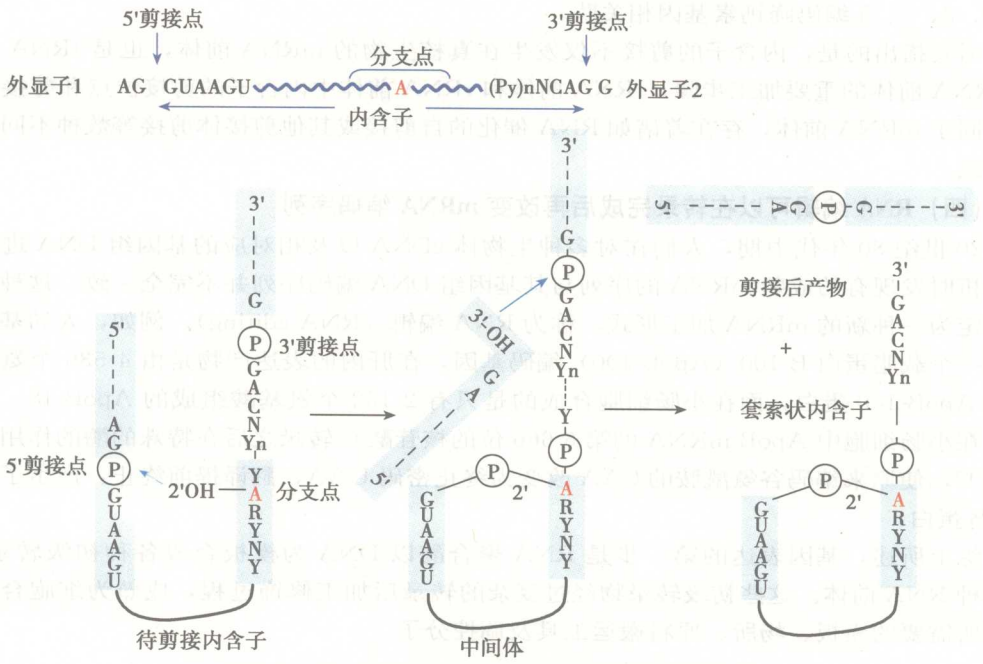


图 5-8　mRNA 内含子剪接特征性序列及套索状中间体
Y：嘌呤核苷酸；R：嘧啶核苷酸；N：嘌呤或嘧啶核苷酸

放出来（图 5-9）。剪切下来的内含子很快被降解。

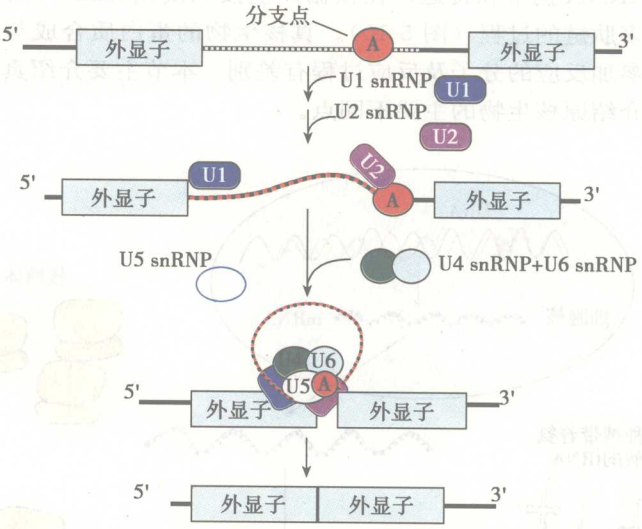


图 5-9　mRNA 剪接体的装配及切除内含子的过程

大部分基因转录时，只产生一种成熟的 mRNA 分子，进而翻译成一种多肽。但是有些基因转录合成的 mRNA 前体可以由于内含子剪接形式的不同而形成不同的成熟 mRNA 分子。这种初级转录物以不同剪接方式选择性地使用外显子、得到不同的 mRNA 分子的现象称为选择性剪接（alternative splicing）。人类基因约有 40%的基因具有选择性剪接的形式，一些基因的 mRNA 前体可能存在多种选择性剪接方式。选择性剪接增加了蛋白质的多态性，例如哺乳动物降钙素基因，利用外显子 1、2、3、4 编码降钙素，而用外显子

1、2、3、5、6 编码降钙素基因相关肽。

需要指出的是，内含子的剪接不仅发生在真核生物的 mRNA 前体，也是 tRNA 前体和 rRNA 前体的重要加工步骤。tRNA 前体和 rRNA 前体中内含子的剪接位点和剪接机制都不同于 mRNA 前体，存在着诸如 RNA 催化的自剪接或其他剪接体剪接等数种不同的剪接方式。

（四）RNA 编辑可以在转录完成后再改变 mRNA 编码序列

20 世纪 80 年代中期，人们在对多种生物体 cDNA 以及相对应的基因组 DNA 进行测序分析时发现有的成熟 mRNA 的序列与其基因组 DNA 编码序列并不完全一致。这种现象被确定为一种新的 mRNA 加工形式，称为 RNA 编辑（RNA editing）。例如，人的基因组只有一个载脂蛋白 B-100（ApoB-100）编码基因，在肝内的表达产物是由 4 536 个氨基酸组成 ApoB-100 蛋白，而在小肠细胞合成的是只有 2 152 个氨基酸组成的 ApoB-48。这是由于在小肠细胞中 ApoB mRNA 的第 6 666 位的核苷酸 C 转录之后在特殊的酶的作用下被改为 U，使原来编码谷氨酰胺的 CAA 改变为终止密码 UAA，翻译提前终止，产生了短的 ApoB 蛋白。

综上所述，基因表达的第一步是 RNA 聚合酶以 DNA 为模板合成各种初级转录物，即各种 RNA 前体。这些初级转录物经过复杂的转录后加工修饰过程，成熟为细胞合成蛋白质所需要的模板、场所、原料搬运工具及调控分子。

第二节　蛋白质的生物合成——翻译

细胞内的蛋白质依据 mRNA 编码信息而合成，故称为翻译。翻译是指组成蛋白质的氨基酸由其特定的 tRNA 携带和转运，在核糖体上按照模板 mRNA 所提供的编码信息合成具有特定序列的多肽链的过程（图 5-10）。真核生物的蛋白质合成与原核生物的蛋白质合成原理相同，但参加反应的分子及反应过程有差别。本节主要介绍真核细胞的蛋白质合成过程，同时兼顾介绍原核生物的主要不同点。

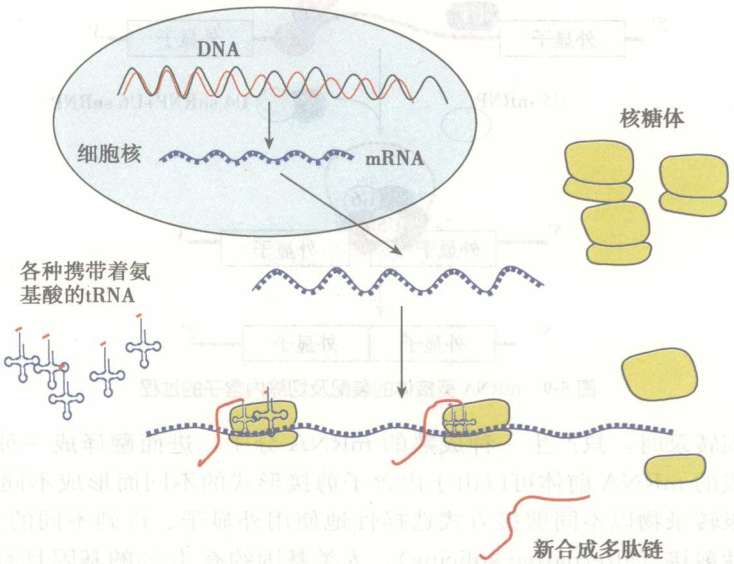


图 5-10　蛋白质生物合成的基本原理

一、蛋白质合成体系

蛋白质生物合成是细胞最为复杂的活动，参与细胞内蛋白质生物合成的物质除原料氨基酸外，还需要 mRNA 作为模板、tRNA 作为特异的氨基酸搬运工具、核糖体作为蛋白质合成的装配场所、有关的酶与蛋白质因子参与反应、ATP 或 GTP 提供能量。

（一）mRNA 是蛋白质合成的信息模板

基因组的遗传信息如何为蛋白质编码这一重要问题由于 mRNA 的发现得以解决。从 DNA 分子中转录而来的 mRNA 分子在细胞质内作为蛋白质合成的模板，mRNA 分子中编码区的核苷酸序列作为遗传密码(genetic codes)，在蛋白质合成过程中被翻译成蛋白质中的氨基酸序列。生物对 mRNA 分子中核苷酸序列的翻译方式以 3 个相邻核苷酸为单位进行，如 AUG 被识别为甲硫氨酸和肽链合成起始信号、UCU 允许丝氨酸进入新生肽链、UGA 将被识别为终止信号。这样串联排列的 3 个核苷酸被称为一个三联体密码子(coden)。4 种构成 mRNA 的核苷酸经排列组合可构成 64 个密码子(表 5-3)，其中 61 个编码 20 种直接在蛋白质合成中使用的氨基酸；另有 3 个不编码任何氨基酸，而作为肽链合成的终止密码子(terminator codon)。

表 5-3 遗传密码

第一个核苷酸（5′端）	第二个核苷酸				第三个核苷酸（3′端）
	U	C	A	G	
U	苯丙氨酸	丝氨酸	酪氨酸	半胱氨酸	U
	苯丙氨酸	丝氨酸	酪氨酸	半胱氨酸	C
	亮氨酸	丝氨酸	终止信号	终止信号	A
	亮氨酸	丝氨酸	终止信号	色氨酸	G
C	亮氨酸	脯氨酸	组氨酸	精氨酸	U
	亮氨酸	脯氨酸	组氨酸	精氨酸	C
	亮氨酸	脯氨酸	谷氨酰胺	精氨酸	A
	亮氨酸	脯氨酸	谷氨酰胺	精氨酸	G
A	异亮氨酸	苏氨酸	天门冬酰胺	丝氨酸	U
	异亮氨酸	苏氨酸	天门冬酰胺	丝氨酸	C
	异亮氨酸	苏氨酸	赖氨酸	精氨酸	A
	*甲硫氨酸	苏氨酸	赖氨酸	精氨酸	G
G	缬氨酸	丙氨酸	天门冬氨酸	甘氨酸	U
	缬氨酸	丙氨酸	天门冬氨酸	甘氨酸	C
	缬氨酸	丙氨酸	谷氨酸	甘氨酸	A
	缬氨酸	丙氨酸	谷氨酸	甘氨酸	G

* 位于 mRNA 起始部位的 AUG 为肽链合成的起始信号。作为起始信号的 AUG 具有特殊性，在原核生物中此种密码代表甲酰甲硫氨酸，在真核生物中代表甲硫氨酸

遗传密码具有以下几个重要特点：

1. 方向性 mRNA 的起动信号到终止信号的排列是有一定方向性的。起动信号总是位于 mRNA 的 5′侧，终止信号总是在 3′侧。mRNA 分子中遗传信息的排列具有方向性（从 5′端至 3′端），决定了翻译过程肽链从 N 端向 C 端合成的方向性。

2. 连续性 mRNA 的密码子之间没有间隔核苷酸。从起始密码子开始，密码子被连续阅读，直至终止密码子出现。这一段序列被称为开放阅读框（open reading frame，ORF）。由于密码子的连续性，在开放阅读框中发生插入或缺失 1 个或 2 个碱基的基因突变，都会引起 mRNA 的阅读框移位（frame shift），使后续的氨基酸序列大部分被改变，

其编码的蛋白质彻底丧失功能；如同时连续插入或缺失 3 个碱基，则只会在蛋白产物中增加 1 个或缺失 1 个氨基酸，但不会导致阅读框移位，对蛋白质的功能影响相对较小。

3. 简并性　64 个密码子中有 61 个编码氨基酸，而氨基酸只有 20 种，因此有的氨基酸可由多个密码子编码，这种现象被称为简并（degeneracy）。例如，UUU 和 UUC 都是苯丙氨酸的密码子，UCU、UCC、UCA、UCG、AGU 和 AGC 都是丝氨酸的密码子。密码子 AUG 具有特殊性，不仅代表甲硫氨酸，如果位于 mRNA 起始部位，它还代表肽链合成的起始密码子（initiator codon）。

4. 摆动性　密码子的翻译通过与 tRNA 的反密码子配对反应而实现。这种配对有时并不严格遵循 Watson-Crick 碱基配对原则，出现摆动（wobble）。此时 mRNA 密码子的第 1 位和第 2 位碱基（5′→3′）与 tRNA 反密码子的第 3 位和第 2 位碱基（5′→3′）之间仍为 Watson-Crick 配对，而反密码子的第 1 位碱基与密码子的第 3 位碱基配对存在碱基配对摆动现象（表 5-4）。

表 5-4　密码子第三位碱基配对的摆动

tRNA 反密码子第一位碱基	密码子第三位碱基	tRNA 反密码子第一位碱基	密码子第三位碱基
C	G	G	C 或 U
A	U	I	U、C 或 A
U	A 或 G		

5. 通用性　从细菌到人类都使用着同一套遗传密码，这一方面为地球上的生物来自同一起源的进化论提供了有力依据，另外使我们有可能利用细菌等生物来制造人类蛋白质。遗传密码的通用性中仍有个别例外，在哺乳动物线粒体内有些密码子编码方式不同于通用遗传密码，如 UAG 不代表终止信号而代表色氨酸，CUA、AUA 编码有所不同，此外终止密码子亦不一样。

（二）各种 tRNA 通过反密码子-密码子识别为 mRNA 模板提供氨基酸原料

在蛋白质合成中使用的 20 种氨基酸各由其特定的 tRNA 负责转运至核糖体。在翻译时，带着不同氨基酸的各种 tRNA 通过其特有的反密码子识别，准确地在核糖体上与 mRNA 上相应的密码子对号入座。细胞内有 40～50 种 tRNA，因此一种氨基酸可由数种 tRNA 携带。

（三）核糖体是肽链“装配厂”

合成肽链所需要的 mRNA 与 tRNA 结合、肽键形成等过程全部是在核糖体上完成的。核糖体类似于一个移动的多肽链“装配厂”，沿着模板 mRNA 链从 5′端向 3′端移动。在此期间，携带着各种氨基酸的 tRNA 分子依据密码子与反密码子配对关系快速进出其中，为延伸的肽链提供氨基酸原料；肽链合成完毕，核糖体立刻离开 mRNA 分子。

在蛋白质合成中除了有 3 种 RNA 以外，还需要由 ATP、GTP 提供能量，Mg^{2+} 离子以及起始因子、延长因子、释放因子和氨基酸-tRNA 合成酶等多种蛋白质的参与。

二、氨基酰 tRNA 合成酶保证氨基酸与 tRNA 的正确结合

参与肽链合成的氨基酸需要活化并与相应 tRNA 结合，该过程是由氨基酰 tRNA 合成酶（aminoacyl-tRNA synthetase）所催化的耗能反应。主要反应步骤包括：①氨基酰 tRNA 合成酶催化 ATP 分解为焦磷酸与 AMP；②AMP、酶、氨基酸三者结合为中间复合体，其中氨基酸的羧基与磷酸腺苷的磷酸以酐键相联，成为活化的氨基酸；③活化氨基酸与 tRNA 分子的 3′-CCA 末端（氨基酸接纳臂）上的腺苷酸（A）的核糖 2′或 3′位游离羟

基以酯键结合，形成相应的氨基酰 tRNA（图 5-11）。

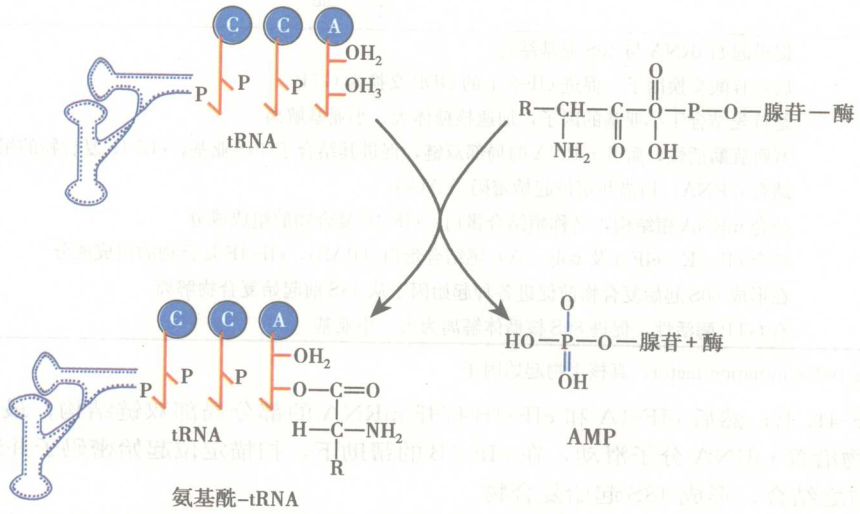


图 5-11　氨基酰 tRNA 的生成

密码子与反密码子间的识别主要由 tRNA 决定而与氨基酸无关，“搭错车”的氨基酸仍将依据 tRNA 的种类进入多肽链导致合成出错。因此氨基酸与 tRNA 连接的准确性是正确合成蛋白质的关键步骤。氨基酸与 tRNA 连接的专一性由氨基酰 tRNA 合成酶保证。20 种氨基酰 tRNA 合成酶都具有高度专一性，既能识别特异的氨基酸，又能辨认应该结合该种氨基酸的 tRNA 分子。氨基酰 tRNA 合成酶还有校对活性，能将错误结合的氨基酸水解释放。氨基酸和 tRNA 结合反应的误差小于 10^{-4}。

三、起始复合物的装配启动新生肽链合成

启动一条多肽链的合成需要将氨基酰 tRNA、mRNA 及核糖体装配成为一个 RNA 和蛋白质构成的复合体。该复合体被称为翻译起始复合物，其装配是一复杂的耗能过程，除核糖体、mRNA、氨基酰 tRNA 外，还需要其他一些蛋白质分子。参与翻译起始复合物装配的蛋白质被称为起始因子（initiation factor，IF）。原核生物只有 3 种起始因子，而真核生物需要至少 10 种以上的起始因子（表 5-5）。

（一）真核生物翻译起始复合物的装配可以分为 3 个阶段

1. 形成 43S 前起始复合物　主要是完成核糖体小亚基与起始 tRNA 的结合。首先 eIF-3 与核糖体 40S 小亚基结合使其保持在与大亚基分离的状态，eIF-6 亦参与大、小亚基核糖体解离；然后在 GTP-eIF-2 帮助下，起始 tRNA 结合到 40S 小亚基上，形成 43S 前起始复合物。

真核生物的起始 tRNA（Met-$tRNA_i^{Met}$）上携带着甲硫氨酸，下标“i”则代表起始（initiation）。这种 tRNA 不同于一般的携带甲硫氨酸的 tRNA（Met-$tRNA^{Met}$），Met-$tRNA_i^{Met}$ 仅识别 mRNA 起始部位的 AUG，而 Met-$tRNA^{Met}$ 仅识别非起始部位的 AUE。原核生物的起始 tRNA 携带的是甲酰化甲硫氨酸（fMet-$tRNA_i^{Met}$）。

2. 形成 48S 前起始复合物　指的是携带有 Met-$tRNA_i^{Met}$ 的 43S 前起始复合物在 mRNA 的翻译起始区域准确就位的过程，该过程在 eIF-4F 复合物（由 eIF-4E、eIFG 和 eIF-4A 组成）和 eIF-4B 的帮助下完成。首先 eIF-4E 结合于 mRNA 5′端的帽结构，而 eIF4G

表 5-5　真核生物翻译起始复合物中的部分起始因子

起始因子	功　能
eIF-2	促进起始 tRNA 与 40S 亚基结合
eIF-2B	属鸟苷酸交换因子，促进 eIF-2 上的 GDP 交换为 GTP
eIF-3	是首先结合于小亚基的因子，加速核糖体大、小亚基解离
eIF-4A	有解旋酶活性；解开 mRNA 的局部双链，促进其结合于 40S 亚基；eIF-4F 复合物的组成成分
eIF-4B	结合 mRNA，扫描并定位起始密码子 AUG
eIF-4E	结合 mRNA 帽结构，又称帽结合蛋白；eIF-4F 复合物的组成成分
eIF-4G	结合 eIF-4E、eIF-3 及 poly（A）尾结合蛋白（PAB），eIF-4F 复合物的组成成分
eIF-5	在形成 80S 起始复合物前促进各种起始因子从 48S 前起始复合物解离
eIF-6	有 GTP 酶活性，促进 80S 核糖体解离为大、小亚基

eIF：eukaryotic initiation factor，真核生物起始因子

结合在 eIF-4E 上；然后 eIF-4A 和 eIF-4B 打开 mRNA 的部分局部双链结构；最后 43S 前起始复合物沿着 mRNA 分子滑动，在 eIF-4B 的帮助下，扫描定位起始密码子并通过 Met-$tRNA_i^{Met}$ 与之结合，形成 48S 起始复合物。

需要指出的是，起始密码子 AUG 同时又为甲硫氨酸编码，但是翻译起始并不发生在 mRNA 分子中的其他甲硫氨酸处。这是由于翻译起始不仅仅依靠 Met-$tRNA_i^{Met}$ 对起始密码子的识别，核糖体中的 rRNA 和蛋白质亦参与对 mRNA 起始密码子周围的序列的识别以决定真正的肽链合成起始点。真核生物的真正的翻译起始密码子常位于一段特定的通用序列 CCRCC AUG G 中（R 为 A 或 G），该序列被称为 Kozac 共有序列（Kozac consensus sequence）。

3. 形成 80S 起始复合物　一旦 48S 复合物定位于起始密码子，eIF-2 上结合的 GTP 即在 eIF-5 的作用下水解为 GDP 后并从 48S 起始复合物中解离，继而导致其他起始因子离开 48S 起始复合物。此时 60S 核糖体大亚基即可结合到 48S 前起始复合物，完成 80S 起始复合物的装配（图 5-12）。

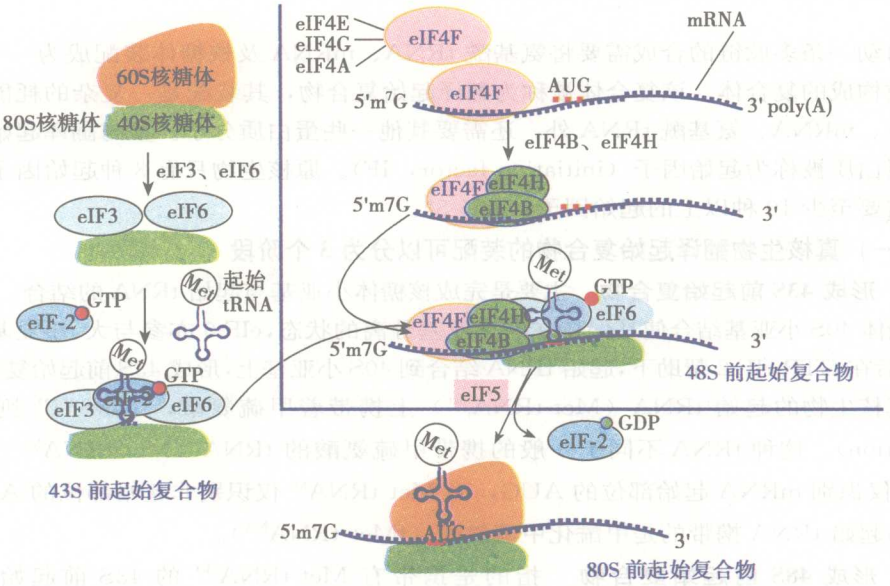


图 5-12　真核生物的翻译起始过程

（二）原核生物蛋白质生物合成起始复合物的特点

原核生物蛋白质生物合成的起始复合物形成过程与真核生物有所不同。主要不同点包括：①起始密码子上游存在一富含嘌呤区（如 AGGA 或 GAGG），称为 SD 序列（Shine-Dalgarno sequence，SD sequence），与核糖体中的 16S rRNA 部分序列互补，两者的结合决定着原核细胞 mRNA 上起始复合物的形成部位；②起始 tRNA 携带的是甲酰甲硫氨酰 $tRNA_i^{Met}$；③起始复合物是 70S；④仅需要 3 种起始因子。其中 IF3 保证核糖体大、小亚基的解离；促进 mRNA 与小亚基的结合，IF1 起辅助作用；IF2 负责识别甲酰甲硫氨酰 $tRNA_i^{Met}$，促进 Met-$tRNA_i^{Met}$ · IF-2 · GTP 三元复合物的形成。

四、新生肽链在核糖体上以连续循环形式进行延伸

肽链在核糖体从 mRNA 的 5′端向 3′端移动过程中依据密码子顺序从 N 端开始向 C 端延长。这是一个在核糖体上重复进行的进位、成肽和转位的循环过程（图 5-13），每完成一个循环，肽链上即可增加一个氨基酸残基。该过程被称为核糖体循环（ribosome cycle）。进位（entrance）或称为注册（registration）指的是一个与 mRNA 密码子相对应的氨基酰-tRNA 进入核糖体；成肽（peptide bond formation）是两个氨基酸间形成肽键的反应；转位（translocation）指核糖体沿着 mRNA 移动一个密码子距离的过程。这一过程除了需要 mRNA、tRNA 和核糖体外，还需要数种延长因子（elongation factor，EF）以及 GTP 等参与。真核生物的 EF 被称为 eEF（eukaryotic elongation factor）。

（一）进位

在 80S 复合物中的核糖体大亚基上有两个氨基酰-tRNA 结合位点。其中一个称为肽酰位或 P 位（peptidyl site，P site），另一个是氨基酰位或 A 位（aminoacyl site，A site）。在起始复合物中，对应于起始密码子的 Met-$tRNA_i^{Met}$ 位于 P 位；此时 A 位是空闲的，并对应着开放阅读框的第二个密码子，配对于此密码子的氨基酰-tRNA 即可进入 A 位。氨基酰-tRNA 进位时需要形成 GTP 复合物，这一三元复合物（氨基酰-tRNA-GTP）的形成在真核生物需要 eEF-1、在原核生物需要 EF-T 参与。

（二）成肽

肽基转移酶催化两个氨基酸间的肽键形成。在起始复合物中，它将 P 位上 tRNA 所携带的甲硫氨酰转移给 A 位上新进入的氨基酰 tRNA，与其所带的氨基酰的氨基结合，形成肽键。二肽形成后，从第三个氨基酸开始，肽基转移酶催化的是 P 位上 tRNA 所连接的肽酰基与 A 位氨基酰间的肽键形成。成肽反应不需要能量供应，也未发现有 EF 的参与。

需要指出的是，肽基转移酶的化学本质不是蛋白质，而是 RNA，因此属于一种核酶。原核生物中大亚基中的 23S rRNA 具有肽基转移酶的活性，在真核生物中，该酶的活性位于大亚基的 28S rRNA 中。

（三）转位

指的是核糖体的移位。成肽反应完成后，核糖体需要向 mRNA 的 3′端移动一个密码子的距离。移位的结果是：①成肽后位于 P 位的空载 tRNA（其携带的氨基酸或肽已经在成肽反应中交给了 A 位上的氨基酸）从核糖体脱落；②成肽后位于 A 位的带有正在合成肽链的 tRNA（肽酰-tRNA）转到了核糖体的 P 位上；③核糖体的 A 位得以空出，且准确定位在 mRNA 的下一个密码子，以接受一个新的对应的氨基酰 tRNA 进位。

转位需要 GTP，此过程需要延长因子的帮助。真核生物的转位过程需要的是延长因子 eEF-2。该延长因子的含量和活性直接影响蛋白质合成速度，因此在细胞适应环境变化

过程中是一个重要的调控靶点。原核生物的转位步骤需要 EF-G，且过程与真核生物稍有不同。在原核细胞，成肽后位于 P 位的空载 tRNA 要先进入一个核糖体上的 tRNA 出口位，或称 E 位（exit site），然后再脱落。

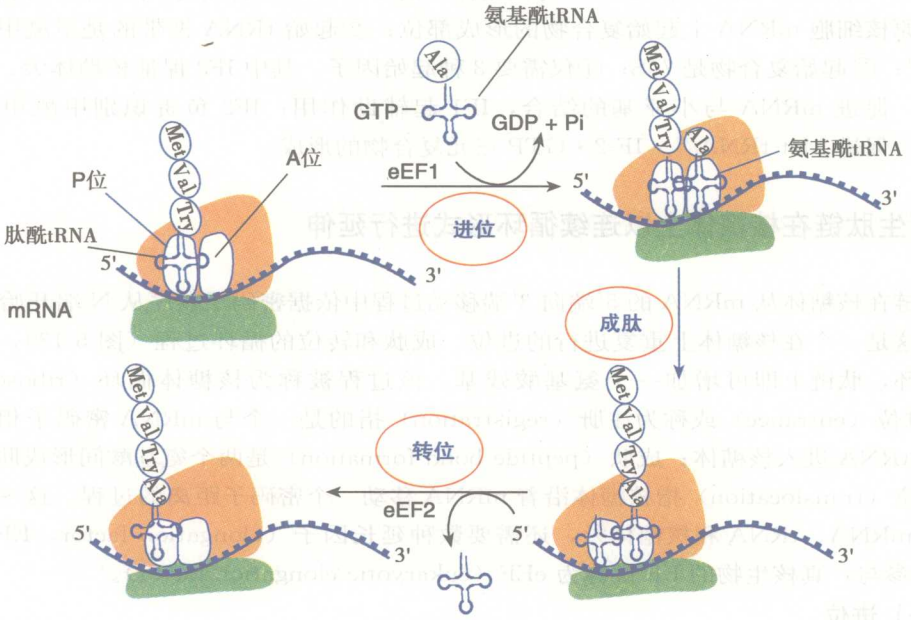


图 5-13　真核生物的核糖体循环

在肽链延长阶段中，每生成一个肽键，都需要直接从 2 分子 GTP（进位与转位各 1 分子）获得能量，即消耗 2 个高能磷酸键化合物；但考虑到氨基酸被活化生成氨基酰 tRNA 时，已消耗了 2 个高能磷酸键，所以在蛋白质合成过程中，每生成一个肽键，实际上共需消耗 4 个高能磷酸键。

五、终止密码子导致肽链合成停止

肽链上每增加一个氨基酸残基，就需要经过一次进位、成肽和转位反应。如此往复，直到核糖体的 A 位对应到了 mRNA 的终止密码子上（UAA，UGA 和 UAG 三者之一），肽链合成即进入终止阶段。在终止阶段，合成完毕的肽链被水解释放，核糖体和 tRNA 从 mRNA 上脱落。这一阶段需要释放因子（release factor，RF）或称终止因子的参与。原核生物有 3 种 RF，其中 RF1 识别终止密码子 UAA 或 UAG，RF2 识别 UAA 或 UGA；RF3 则与 GTP 结合并使其水解，协助 RF1 与 RF2 与核糖体结合。多肽链在肽酰-tRNA 之间的酯键水解后从核糖体上释放，该反应仍然是由催化成肽反应的肽基转移酶完成，条件是有 RF 和 GTP 存在。

真核生物仅有 eRF 一种释放因子，所有 3 种终止密码子均可被 eRF 识别。真核生物中肽链合成完成后的水解释放过程尚未完全了解，推测可能有其他的蛋白分子参与这一过程。

在一条 mRNA 上可以同时附着多个核糖体，形成多聚核糖体，同时合成多条同样的多肽链，保证蛋白质生物合成的高效率。

六、新生肽链经加工后成为具有功能的蛋白质

新生肽链并不具有生物活性，它们必须正确折叠形成具有生物活性的三维空间结构，有的要形成必需的二硫键，有的需经过亚基的聚合形成具有四级结构的蛋白质。除此以外，许多蛋白质在翻译后还需经过蛋白水解作用切除一些肽段或氨基酸，或对某些氨基酸残基的侧链基团进行化学修饰等处理后才能成为有活性的成熟蛋白质。

（一）肽链的正确折叠是形成功能蛋白质的基础

细胞内的新生多肽链在翻译进行的过程中已经开始折叠，折叠方式由其氨基酸的排列顺序所决定（一级结构决定空间结构）。细胞中存在一类专门协助多肽链折叠的蛋白质，它们被称为分子伴侣（molecular chaperone）。分子伴侣广泛存在于原核细胞和真核细胞。在真核细胞中，又广泛分布在细胞质、线粒体、内质网及胞核等部位。分子伴侣的可能作用有：①封闭待折叠多肽链中暴露的疏水区段；②创造一个隔离的环境，蛋白质可互不干扰进行折叠；③促进折叠；④在某些应激状态下，使已折叠的蛋白质去折叠。目前已知的分子伴侣有热休克蛋白（heat shock protein，Hsp）家族中的 Hsp70、Hsp 60、Hsp 90 以及蛋白质二硫键异构酶、肽酰脯氨酸异构酶等。二硫键异构酶催化多肽链的链内与链间二硫键的形成，对维持蛋白质立体结构具有重要作用。

（二）肽链需要进行特定的加工和修饰

1. 某些肽段或氨基酸残基的切除 新生多肽链的有限水解是一种最常见的翻译后加工形式，几乎所有成熟的多肽链都要经过这种形式的加工。有许多参与机体不同生理过程的蛋白质，其初始形式都是不具有活性的前体，然后通过有限的蛋白水解作用，去除某些肽段后，成为有活性的蛋白质分子。新生肽链的 N 端的甲硫氨酸残基，在肽链离开核糖体后，即由特异的蛋白水解酶切除。分泌性蛋白和跨膜蛋白的前体的 N 端都有一含 13～36 个氨基酸残基（以疏水氨基酸残基为主）的肽段——信号肽（signal peptide），这些信号肽在蛋白质成熟过程中需要被切除。有的多肽链经水解可以产生数个不同的多肽链，例如，由垂体前叶和中叶合成的阿片黑素皮质素原（pro-opio-melanocortin，POMC）就可以加工成 7 种不同的多肽激素。

2. 氨基酸残基的共价修饰 多肽链翻译后加工的另一种主要形式是肽链的某些氨基酸残基侧链基团的共价化学修饰，这些修饰常发生在重要的功能基团。除了丙氨酸、甘氨酸、亮氨酸、异亮氨酸、甲硫氨酸和缬氨酸外，其他氨基酸残基的侧链都可被修饰。常见的修饰有乙酰化、甲基化、糖基化、羟化、核苷酸化、磷酸化、ADP-核糖化等。蛋白质化学修饰是蛋白质活性调节的主要方式，也是影响蛋白质分子之间的相互作用的重要方式，因此在细胞增殖、代谢调节、基因表达调控等重要细胞活动中都具有十分重要的作用。

（三）亚基组装形成具有四级结构的蛋白质

生物体内的功能蛋白质许多是由一条以上肽链构成的蛋白质。多聚体的肽链之间通过非共价键维持一定空间构象，有些还需与辅基聚合才能形成具有活性的蛋白质。由一条以上肽链构成的蛋白质的各个亚单位相互聚合时所需要的信息蕴藏在肽链的氨基酸序列之中，而且这种聚合过程往往又有一定顺序，前一步骤常可促进后一聚合步骤的进行。成人血红蛋白由两条 α 链、两条 β 链及 4 个血红素分子组成。合成好的 α 链从核糖体上自行释放，与尚未从核糖体释放的 β 链相结合，与 β 链同时离开核糖体上，形成游离的 α、β 二聚

体。此二聚体再与线粒体内生成的两个血红素相结合，最后才形成一个由 4 条肽链和 4 个血红素构成的有功能的血红蛋白分子。

（四）蛋白质合成后需要转运到特定部位才能发挥其生物功能

为参加生命活动，在细胞质合成的蛋白质还必须准确地到达其发挥功能的亚细胞区域或分泌到细胞外。蛋白质的亚细胞定位信息存在于其自身结构中，如位于细胞核内的蛋白质结构中有特定氨基酸序列构成的核定位信号，而位于细胞质膜的蛋白和分泌蛋白有信号肽等，因而现代生物信息学可以从基因的结构推测其编码的蛋白质在细胞内的可能位置。蛋白质的转运方式可分为两大类：一种是细胞内的蛋白质合成与转运同时发生，即翻译转运同步；另一种是蛋白质从核糖体上释放后才发生转运，属翻译后转运。

（五）分泌性蛋白质的合成与转运同步进行

分泌性蛋白多采用翻译转运同步机制。它们的 N 端都有信号肽结构，可以被细胞质中的信号识别颗粒（signal recognition particle，SRP）所识别。SRP 是由 7SL-RNA 和 6 种不同的多肽链组成的 RNA-蛋白质复合体。含有信号肽的多肽的翻译转运机制和步骤包括：①在核糖体上合成时，由于信号肽部分位于 N 端故首先被合成出来，并被 SRP 所捕捉，SRP 随即结合到该核糖体上；②内质网膜上有 SRP 的受体，因此 SRP-核糖体复合体可以被引导到内质网膜上；③在内质网膜上，通过两种内质网膜蛋白使正在合成的肽链穿过内质网膜孔进入内质网；④SRP 脱离信号肽和核糖体，肽链继续延伸直至完成；⑤信号肽在内质网内被信号肽酶（signal peptidase）切除；⑥肽链在内质网中折叠形成最终构象，最后又经高尔基体转运到胞外。

（六）胞浆蛋白质是合成后被转运的

另外一些蛋白质是在肽链翻译完成以后被转运到其功能区域的。例如，真核细胞线粒体中的绝大部分蛋白质是由核染色体基因组 DNA 编码、在细胞质合成的。这些蛋白质需经跨膜转运方可进入线粒体的各部分。可通过线粒体膜的蛋白质在转运前大多数是以前体形式存在，在它的 N 端有一段前导肽，当前体蛋白通过线粒体膜时前导肽被多肽酶水解，前体蛋白成为成熟蛋白质。蛋白质通过线粒体内膜的转运需要消耗能量，而外膜上可能有受体参与转运作用。

七、真核生物蛋白质合成的其他特点

真核生物蛋白质合成的原理和基本反应过程与原核生物相同，但是两者的具体反应步骤和方式亦有许多差异。除了上述已经提及的一些反应步骤和方式上的不同点以外，真核生物的蛋白质生物合成还有以下几个重要的特点：

1. 真核生物的蛋白质合成与 mRNA 的转录分开进行，而原核生物中翻译与转录过程偶联在一起同时进行。真核生物的 mRNA 的转录与加工在细胞核内完成后需要转运到细胞质，蛋白质的生物合成在细胞质进行。

2. 真核生物中一条 mRNA 模板只含有一个翻译起始点和一个终止点，因而只能翻译出一条多肽链。这种 mRNA 被称为单顺反子（monocistron）；而原核生物的大部分 mRNA 是多顺反子，其中可有两个以上的翻译起始点和终止点，因而同时翻译出几种不同的多肽链。

3. 真核生物蛋白合成体系中的 RNA 和蛋白质在结构上与原核生物有一定差异，因而对于一些小分子蛋白质合成抑制剂的敏感性不同。临床上某些用于抗菌治疗的抗生素可以

抑制原核生物的蛋白质生物合成，但是对真核生物并无影响，因而不必顾虑相应的副作用。真核细胞蛋白质合成抑制药物可能用于某些肿瘤的治疗，但副作用较大。另外真核生物的线粒体有独立的蛋白质生物合成体系，且类似于原核生物，这可能是某些抗菌药物仍可产生副作用的原因。

（赵涵芳）

第六章　基因表达调控

关键知识点

- 生物体基因组中基因的表达根据生长、分化和发育的需要，或者随着环境的变化而具有时间特异性和空间特异性。基因表达的种类和强度决定了细胞的分化状态和功能。
- 基因表达在转录水平的调控主要取决于基因组的顺式作用元件与反式作用因子的相互作用，是一个蛋白质与 DNA 以及蛋白质与蛋白质相互作用的复杂过程。基因表达调控贯穿于基因表达的转录和翻译全过程。
- 操纵子是原核生物基因表达调控的基本单位。乳糖操纵子由于阻遏蛋白结合于操纵序列而关闭，该阻遏能力可因诱导剂别乳糖的结合而失去，故别乳糖可以诱导乳糖操纵子开放。正调控蛋白 CAP 的结合可进一步提高乳糖操纵子的转录效率。色氨酸操纵子存在一特殊的衰减调控作用，可使 RNA 的转录提前终止。
- 真核生物基因表达可在多层次上进行调控。组蛋白的共价修饰和 DNA 甲基化引起的染色质结构改变可以影响基因的转录速度。各种转录因子是转录水平的主要调控分子，其结构特征是有 DNA 结合结构域和转录调节结构域。通用转录因子是各组织细胞所共有，为 RNA 聚合酶结合于启动子所必需；组织特异性转录因子的种类、含量和活性在不同组织细胞、不同功能阶段受到环境信号控制，是形成基因表达的时空特异性的主要因素。真核生物 RNA 转录后的加工修饰复杂，可受很多因素调节。翻译起始蛋白的活性、mRNA 分子的寿命、mRNA 的 5′-非翻译区以及 3′-非翻译区结构都会影响细胞蛋白质合成的速度。

关键词： 顺式作用元件；反式作用因子；可诱导型操纵子；操纵序列；阻遏蛋白；转录因子；增强子；miRNA

基因表达调控（gene expression regulation）是指生物体通过特定的蛋白质与 DNA、蛋白质与蛋白质之间的相互作用来控制基因是否表达，或调节表达产物的多少以满足生物体的自身需求以及适应环境变化的过程。这些调节过程受到严格而有序的控制。一般而言，生物物种的级别越高，调节基因表达的过程也就越复杂和越精细。

第一节　基因表达调控的基本规律

原核生物体系和真核生物体系在基因组结构以及细胞结构上的差异使得它们的基因表达方式有所不同。原核细胞没有细胞核，遗传信息的转录和翻译发生在同一空间，并以偶联的方式进行。真核细胞具有细胞核，使得转录和翻译不仅具有空间分布的特征，而且还有时间上的先后顺序。尽管如此，原核生物体系和真核生物体系在基因表达调控上都遵循一些共同的基本规律。

（一）基因表达具有时空特异性

虽然同一生物体的各种细胞含有完全相同的基因组，但是它们在细胞内并非同时表达，而是根据生长、分化和发育等功能的需要，随着环境的变化，按照一定的时间顺序先后表达。这种性质被称为基因表达的时间特异性（temporal specificity），或阶段特异性（stage specificity）。与此相类似，同一基因在不同的组织或器官中的表达也不一样。不仅

基因表达的种类不同，而且基因表达的水平也不相同，这就是真核细胞中基因表达的空间特异性（spatial specificity），或组织特异性（tissue specificity）。同一个体内的不同器官、组织、细胞的差异性的基础是特异的基因表达或称为差异基因表达（differential gene expression）。细胞的基因表达谱（gene expression profile），即基因表达的种类和强度决定了细胞的分化状态和功能。换句话讲，在个体内决定细胞类型的不是基因本身，而是基因表达模式（gene expression pattern）。

（二）诱导表达和阻遏表达是基因表达调控的普遍方式

有些基因参与生命的全过程，因此必须在一个生物体的所有细胞中持续地表达。这样的基因被称为管家基因（housekeeping gene）。管家基因的表达只与启动子和 RNA 聚合酶有关，基本上不受环境因素和其他因素的影响。这样的基因表达方式称为组成性表达（constitutive expression）。与此相对应的是一类受到调控的基因表达。在特定的信号刺激下，有些基因表现出开放性或增强性的表达，称其为诱导（induction）；而另一些则表现出关闭性或抑制性的表达，称其为阻遏（repression）。这些基因分别被称为可诱导基因（inducible gene）和可阻遏基因（repressible gene）。诱导性和阻遏性基因表达是生物体为适应外界环境的改变而做出的两种应答方式。

（三）基因表达受顺式作用元件和反式作用因子共同调节

一个生物体的基因组中既有携带遗传信息的结构基因，也有能够影响结构基因表达的调节序列。一般说来，调节序列与被调控的结构基因位于同一条 DNA 链上，即顺式作用元件。另外一些调节序列（基因）远离被调控的结构基因，实际上是其他分子的编码基因，只能通过其表达产物来发挥作用的。这样的调节基因不仅能对处于同一条 DNA 链上的结构基因的表达进行调控，而且还能对不在一条 DNA 链上的结构基因的表达起到同样的作用。因此，这些调节基因的产物被称为反式作用因子，亦称为转录因子。这些反式作用因子以特定的方式识别和结合在顺式作用元件上，实施精确的基因表达调控。

（四）蛋白质-DNA 以及蛋白质-蛋白质的相互作用是基因表达调控的分子基础

作为反式作用因子的调节蛋白具有特定的空间结构，通过特异性地识别某些 DNA 序列与顺式作用元件发生相互作用。例如，DNA 双螺旋结构的大沟是调节蛋白最容易与 DNA 序列发生相互作用的部位。真核生物体系基因组结构比较复杂，使得有些调节蛋白不能够直接与 DNA 相互作用，而是首先形成蛋白质-蛋白质的复合物，然后再与 DNA 结合参与基因表达的调控。

（五）基因表达调控是多层次的复杂调节

无论是原核生物还是真核生物，基因表达的调控都体现在基因表达的全过程中，即在 RNA 转录合成和蛋白质翻译两个阶段都有控制其表达量的机制，因此基因表达的调控是多层次的复杂过程。真核生物由于基因组的庞大、DNA 高级结构的复杂、转录产物复杂的加工修饰、转录与翻译的分隔进行、mRNA 结构的特殊性等诸多因素，基因表达的调控过程较原核生物更要复杂许多（图 6-1）。该过程包括了染色质激活、转录起始、转录后修饰、转录产物的细胞内转运、翻译起始、翻译后修饰等多个步骤。在上述过程的每一个环节都可以对基因表达进行干预，从而使得基因表达调控呈现出多层次和综合协调的特点。但是，转录起始的调控是基因表达调控的较为关键的环节。

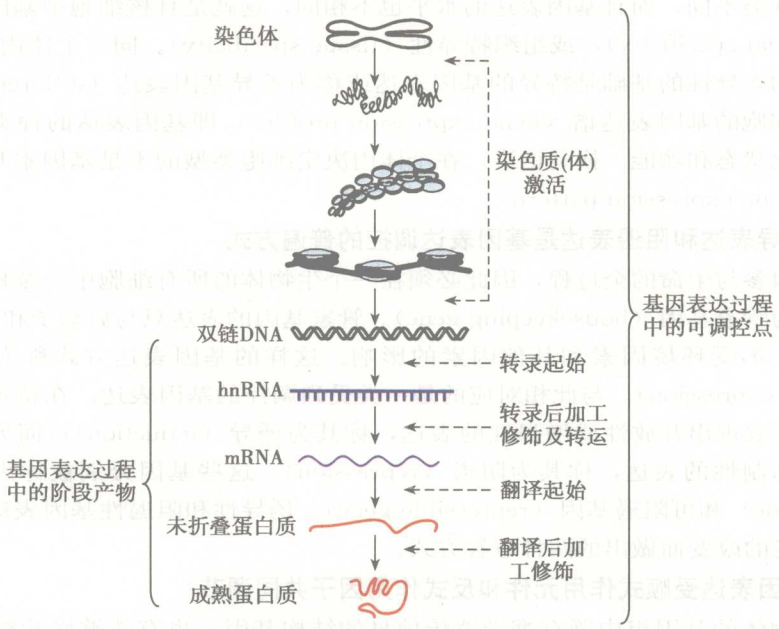


图 6-1　基因表达的多层次复杂调控

第二节　原核生物的基因表达调控

原核生物基因组是一具有超螺旋结构的闭合环状 DNA 分子。原核生物的细胞结构也比较简单，它的基因组的转录和翻译可以在同一空间内完成，并且时间上的差异不大。在转录过程终止之前，mRNA 就已经结合在核糖体上，开始了蛋白质的生物合成。

一、转录水平的调控是原核生物基因表达的关键环节

前已述及，大肠杆菌的 RNA 聚合酶是 σ 亚基（或称 σ 因子）和核心酶构成，其中 σ 亚基的作用是识别和结合在 DNA 模板上的启动子序列，启动转录过程。原核生物在转录水平的调控主要取决于转录起始速度，即主要调节的是转录起始复合物形成的速度。

（一）操纵子是原核生物基因表达调控的基本单元

1961 年，F. Jacob 和 J. L. Monod 二人提出了原核生物基因表达调控的操纵子模型，成功地解释了大肠杆菌在不同环境下利用不同糖源的调节机制。操纵子模型在原核生物基因表达调控中具有普遍性。

操纵子是原核生物基因组构的基本单位，也是基本转录单位，由结构基因和调控序列组成（见第二章）。结构基因编码原核生物代谢所需要的酶，而调控序列包括启动子、操纵序列以及一定距离外的调节基因。操纵序列是可以被调控蛋白特异性结合的 DNA 序列，位于启动子和结构基因之间，与上游的启动子和下游的结构基因可有部分重叠。这种重叠体现了 RNA 聚合酶和调节蛋白与操纵序列结合的竞争特性。

调节基因（regulatory gene）编码能够与操纵序列结合的调控蛋白。通常情况下，调节基因编码的蛋白可以阻遏结构基因的表达，故又称为阻遏蛋白（repressor）。调控蛋白与操纵序列结合后能够减弱或阻止基因转录，其介导的调控方式称为负调控（negative

regulation），是原核生物中较为普遍的调节方式；反之，如果调控蛋白与操纵序列结合后能增强或起动基因转录，则称为正调控（positive regulation）。诱导和阻遏发生的机制是某些特殊的物质能与调控蛋白结合，使调控蛋白的空间构象发生变化，从而改变了对基因转录的影响。凡是能够诱导基因表达的分子称为诱导剂（inducer），而凡是能够阻遏基因表达的分子称为阻遏剂。

（二）乳糖操纵子模型诠释了原核生物转录起始诱导的基本机制

大肠杆菌的乳糖代谢酶的基因表达特点是：在环境中没有乳糖时，这些基因处于关闭状态；只有当环境中有乳糖时，这些基因才被诱导开放，合成代谢乳糖所需要的酶。乳糖操纵子（*lac* operon）是最早发现的原核生物转录调控模式。乳糖操纵子的结构基因 *lacZ*，*lacY* 和 *lacA* 分别编码 β-半乳糖苷酶、通透酶和半乳糖苷转乙酰基酶，它们都是细胞利用乳糖所必需的酶。

乳糖操纵子在不含乳糖的环境中处于关闭状态是由于阻遏蛋白结合在操纵序列（*lacO*）上阻碍了起始复合物形成所致（图 6-2A）。阻遏蛋白编码基因 *lacI* 具有组成性表达特征，其表达产物形成四聚体后与 *lacO* 结合，以此阻止了结合在启动子 P_{lac} 上的 RNA 聚合酶向下游移动，关闭了转录（图 6-2B）。这种阻遏机制保证了在通常情况下乳糖操纵子仅处于基础表达状态，即只有 1/100 的基因得到了表达。

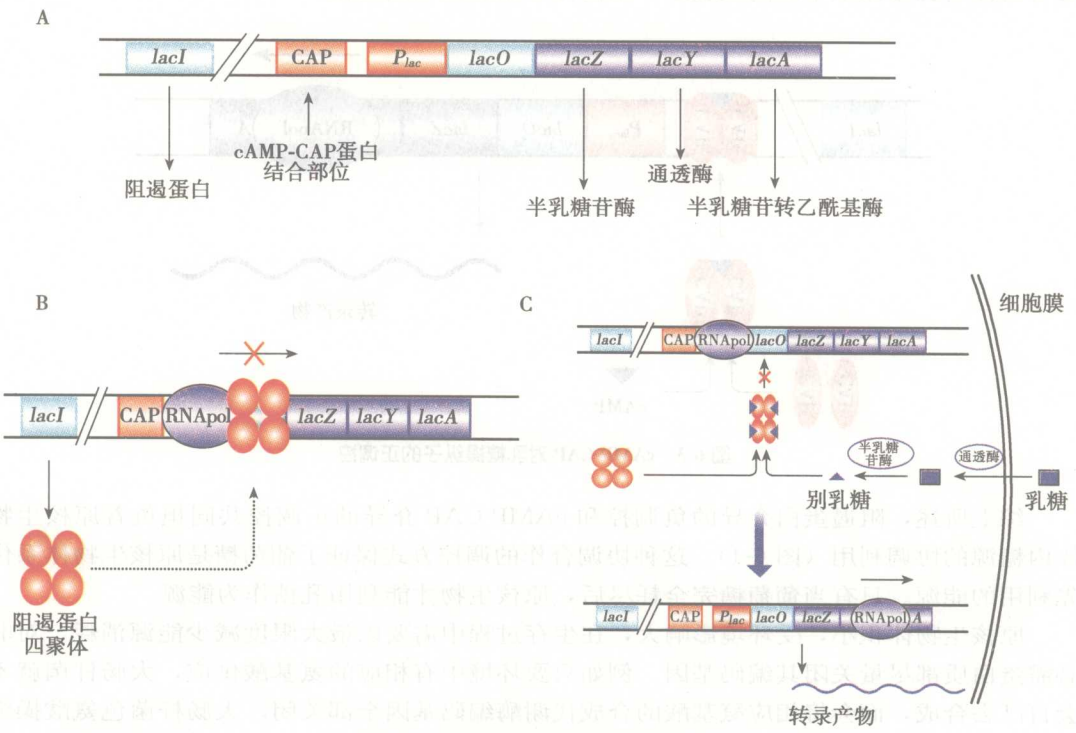


图 6-2　乳糖操纵子的关闭和诱导开放

A. 乳糖操纵子的基本结构；B. 不含乳糖时，结合于操纵序列的阻遏蛋白关闭乳糖操纵子的基因表达；C. 有乳糖时，乳糖进入细胞，生成别乳糖，结合阻遏蛋白，使其发生变构而失去与 *lacO* 结合的能力，从而使得 RNA 聚合酶就可以顺利地通过 *lacO* 区域而起始转录

当环境中出现乳糖时，乳糖在通透酶的作用下进入细胞内，并在 β-半乳糖苷酶的作用下生成葡萄糖和半乳糖，少部分葡萄糖和半乳糖经 β-半乳糖苷酶催化可生成别乳糖。别乳糖作为诱导物，可以结合阻遏蛋白，使其失去与 *lacO* 结合的能力。这样 RNA 聚合酶可以

有效地启动转录，表达出细胞利用乳糖所需要的 3 种酶（图 6-2C）。在基因工程中可以使用乳糖操纵子的启动子来控制所携带基因的表达，此时可用异丙基硫代半乳糖苷（isopropylthiogalactoside，IPTG）来诱导乳糖操纵子的开放。由于别乳糖可被 β-半乳糖苷酶水解重新生成葡萄糖和半乳糖，而 IPTG 则不受 β-半乳糖苷酶的作用，因此它是比别乳糖更强的诱导物。

（三）乳糖操纵子的高效表达需要另外的正调控方式

大肠杆菌在不同环境下的糖源利用还有一特点，就是在同时存在乳糖和葡萄糖的情况下优先利用葡萄糖，此时乳糖操纵子虽然处于开放状态，但是表达效率不高。这是因为乳糖操纵子启动子属于弱启动子，即使在 *lacO* 开放的条件下，仍需要另一个正调控方式来提高转录启动效率。这种正调控方式如图 6-3 所示：在 P_{lac} 上游－60 bp 区有一个分解代谢物基因活化蛋白（catabolite gene activation protein，CAP）的结合位点。CAP 与环腺苷酸（cAMP）结合后形成二聚体，结合在 CAP 结合位点上，由此提高了乳糖操纵子的活性，大大增强了 P_{lac} 的转录起始。细胞中 cAMP 的浓度受葡萄糖浓度控制。当大肠杆菌处在葡萄糖富有环境中时，葡萄糖的分解抑制了腺苷酸环化酶的活性，减少 cAMP 的生成，使得细胞内 cAMP-CAP 复合物的浓度不足以激活 P_{lac}，此时即使存在乳糖的诱导作用，但是转录效率不高。只有在葡萄糖耗尽时，升高的 cAMP 浓度才能形成足够的 cAMP-CAP 复合物促进乳糖操纵子的高效表达，细菌方开始主要利用乳糖。

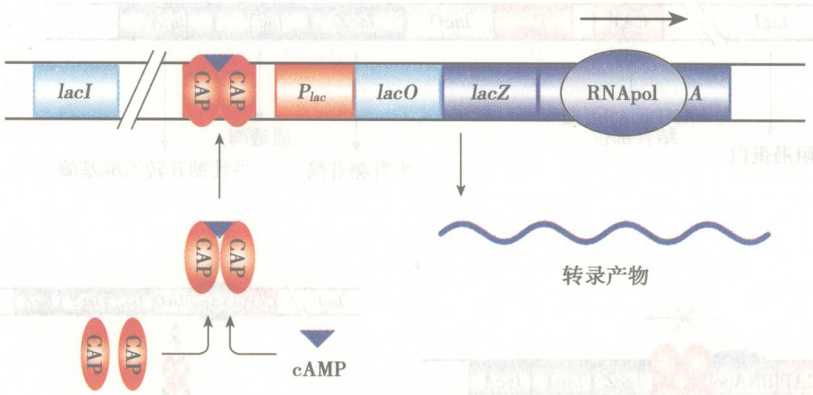


图 6-3　cAMP-CAP 对乳糖操纵子的正调控

综上所述，阻遏蛋白介导的负调控和 cAMP-CAP 介导的正调控共同担负着原核生物体内糖源的协调利用（图 6-4）。这种协调合作的调控方式保证了葡萄糖是原核生物体系优先利用的能源。只有当葡萄糖完全耗尽后，原核生物才能利用乳糖作为能源。

原核生物体积小，受环境影响大，在生存过程中需要以最大限度减少能源消耗，对非必需蛋白质都尽量关闭其编码基因。例如只要环境中有相应的氨基酸供应，大肠杆菌就不会自己去合成，而会将相应氨基酸的合成代谢酶编码基因全部关闭。大肠杆菌色氨酸操纵子（*trp* operon）的结构基因有 *trpE*、*trpD*、*trpC*、*trpB* 和 *trpA*，分别编码合成色氨酸所需的 5 种酶。上游的调控序列由启动子和操纵元件组成。调节基因 *trpR* 亦编码一种阻遏蛋白。在色氨酸操纵子的结构基因和调控序列之间，存在一段长度为 162 bp 的前导序列基因 *trpL*（图 6-5A）。色氨酸操纵子是一个阻遏操纵子。在细胞内无色氨酸时，阻遏蛋白不能与操纵序列结合，因此色氨酸操纵子处于开放状态，结构基因得以表达。当细胞内色氨酸的浓度较高时，色氨酸作为辅阻遏物与阻遏蛋白形成复合物并结合到操纵序列上，关闭色氨酸操纵子（图 6-5B），停止表达用于合成色氨酸的各种酶。

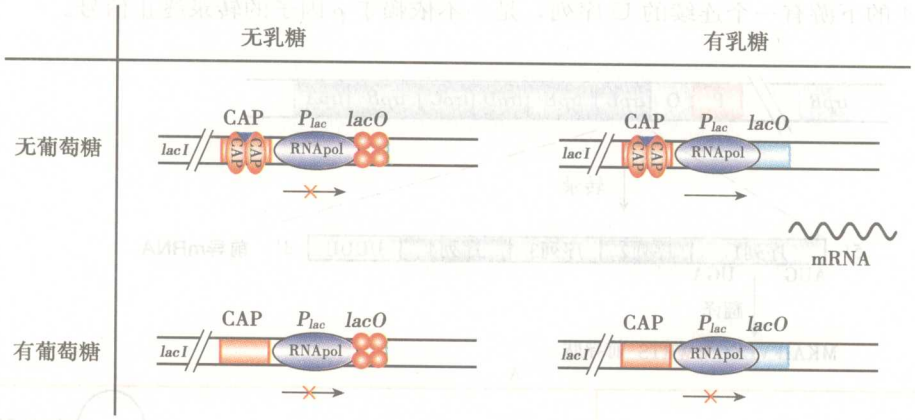


图 6-4　乳糖操纵子的协调控制

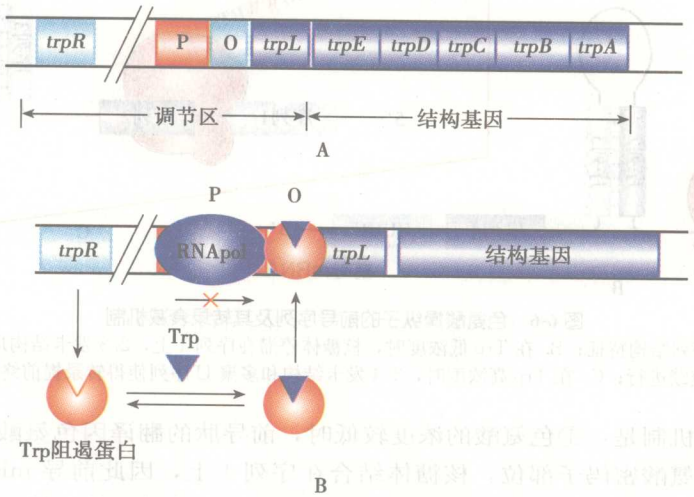


图 6-5　色氨酸操纵子的结构及其关闭机制

A. 色氨酸操纵子的基本结构；B. 高浓度的色氨酸作为阻遏剂与阻遏蛋白结合，并结合在操纵序列上，关闭色氨酸操纵子的转录

二、翻译水平的调控是对转录调控的补充

原核生物中不同的操纵子有不同的调控方式，除了在转录水平的调控以外，在蛋白质翻译水平亦有调节，保证了基因表达调节的精确和有效。

（一）转录与翻译的偶联调控提高了基因表达调控的有效性

上述的色氨酸操纵子的有效关闭还有一种属于促进已经开始转录的 mRNA 合成终止的方式来进一步加强，称为转录衰减（attenuation）。这种作用是利用原核生物中转录与翻译过程偶联进行，翻译时先合成的一段前导序列 L 来实现的。前导序列的结构特点及其发挥衰减作用的机制如图 6-6 所示。前导序列 L 的结构特点是：①它可以转录生成一段长度为 162 bp、内含 4 个特殊短序列的前导 mRNA；②其中序列 1 有独立的起始和终止密码子，可翻译成为一个有 14 个氨基酸残基的前导肽，它的第 10 位和第 11 位都是色氨酸残基；③序列 1 和序列 2 间、序列 2 和序列 3 间、序列 3 和序列 4 间存在一些互补序列，分别都可以形成发夹结构。形成发卡结构的能力依次是 1/2 发夹＞2/3 发夹＞3/4 发夹；

④序列 4 的下游有一个连续的 U 序列，是一不依赖于 ρ 因子的转录终止信号。

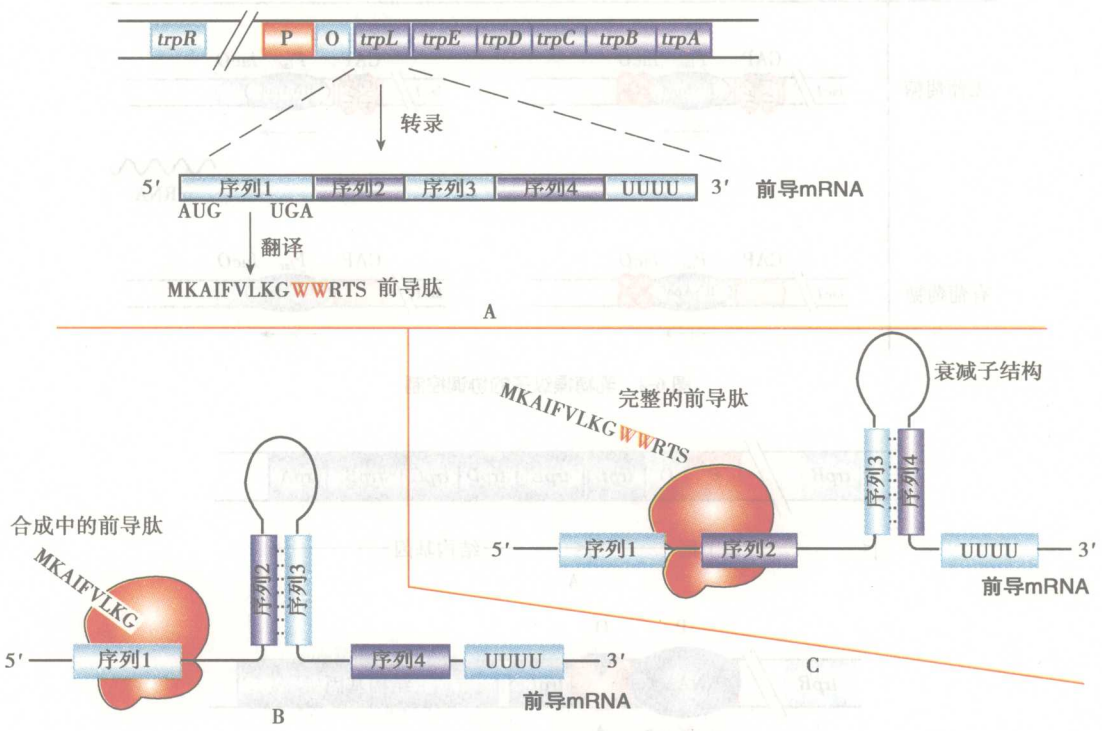


图 6-6　色氨酸操纵子的前导序列及其转录衰减机制

A. 前导序列结构特征；B. 在 Trp 低浓度时，核糖体停滞在序列 1 上，2/3 发卡结构形成，转录继续进行；C. 在 Trp 高浓度时，3/4 发卡结构和多聚 U 序列使得转录提前终止

转录衰减的机制是：①色氨酸的浓度较低时，前导肽的翻译因色氨酸量的不足而停滞在第 10/11 的色氨酸密码子部位，核糖体结合在序列 1 上，因此前导 mRNA 倾向于形成 2/3 发夹结构，转录继续进行；②色氨酸的浓度较高时，前导肽的翻译顺利完成，核糖体可以前进到序列 2，因此发夹结构在序列 3 和序列 4 形成，连同其下游的多聚 U 使得转录中途终止，表现出转录的衰减。原核生物这种在色氨酸低浓度时通过阻遏作用和转录衰减机制共同关闭基因表达的方式，保证了营养物质和能量的合理利用。

前导序列起到了随色氨酸浓度升高而降低转录的作用，故将这段序列称为衰减子（attenuator）。在 *trp* 操纵子中，阻遏蛋白对结构基因转录的负调控起到粗调的作用，而衰减子起到精调的作用。这是由于前导肽中含有两个相连的色氨酸残基所造成的。细菌中其他氨基酸合成系统的操纵子（如 *pheA*、*his*、*leu*、*thr* 等）中也有类似的衰减调控机制。

（二）SD 序列影响翻译起始速度

在多顺反子 mRNA 中，每一个蛋白编码区都是以起始密码子 AUG 开始。在 AUG 上游 8～13 核苷酸位置上的 SD 序列有助于 mRNA 在核糖体的小亚基上准确定位。因此，SD 序列特征、SD 序列与起始密码子之间的距离，以及 mRNA 二级结构对 SD 序列的屏蔽作用等都会影响到翻译起始的效率。

（三）翻译阻遏利用蛋白质与自身 mRNA 的结合实现翻译起始的调控

翻译起始与转录起始相类似，也受调节蛋白的作用，但与转录不同，RNA 在翻译起始过程中有重要的作用。编码区的起始点可与调节分子（蛋白质或 RNA）直接或间接地结合来决定翻译起始。在此调控机制中，调节蛋白可以结合到起始密码子上，阻断与核糖

体的结合。例如 S8 是组成核糖体小亚基的一个蛋白质，可以与 16S rRNA 的茎环结构结合；L5 是组成核糖体大亚基的一个蛋白质，它的 mRNA 的 5′末端也能形成一个与 16S rRNA 的茎环结构相类似的结构。因此 S8 也能与 L5 的 mRNA 结合。当 16S rRNA 含量充足时，可以与所有的 S8 蛋白结合，不影响 L5 蛋白的合成；而当 16S rRNA 含量不足时，多余的 S8 则与 L5 mRNA 结合，阻遏 L5 蛋白质的合成，防止 L5 合成过量。

此外，在一些细菌和病毒中还存在一类调节基因，能够转录产生反义 RNA。反义 RNA 能够与特定的靶 mRNA 序列互补配对结合，形成一个双链区，从而阻断 mRNA 的翻译起始。反义 RNA 作用具有非常重要的理论意义和实际意义。

（四）mRNA 密码子的编码频率影响翻译的延伸速度

遗传密码表显示，除色氨酸和甲硫氨酸外，其他的氨基酸都有 2 个或 2 个以上的遗传密码子。有些是使用频率较高的常用密码子，而有些则是使用频率较低的稀有密码子。当基因中的密码子是常用密码子时，mRNA 的翻译速度快，反之，mRNA 的翻译速度慢。大肠杆菌 *dnaG* 基因是引物酶的编码基因，含有较多的稀有密码子，使得 mRNA 的翻译速度缓慢，防止引物酶合成过多。

第三节　真核生物的基因表达调控

多细胞真核生物的基因表达调控比原核生物复杂得多，这是因为：①真核基因组比原核基因组大得多；②原核基因组的大部分序列都为编码基因，而哺乳类基因组中只有 10% 的序列编码蛋白质、rRNA、tRNA 等，其余 90% 的序列，包括大量的重复序列功能至今还不清楚，可能参与调控；③真核生物编码蛋白质的基因是不连续的，转录后需要剪接去除内含子，这就增加了基因表达调控的层次；④原核生物的基因组织在操纵子中，多顺反子 mRNA 使得几个功能相关的基因自然协调控制；而真核生物则是一个结构基因转录生成一条 mRNA，即 mRNA 是单顺反子，许多功能相关的蛋白，即使是一种蛋白的不同亚基也将涉及多个基因的协调表达；⑤真核生物 DNA 在细胞核内与多种蛋白质结合构成染色质，这种复杂的结构直接影响着基因表达；⑥真核生物的遗传信息不仅存在于核 DNA 上，还存在线粒体 DNA 上，调控既相互独立而又需要协调。

由此可见，真核生物的基因表达调控比原核生物复杂得多。由于细胞核的存在，转录和翻译过程既有空间上的不同又有时间上的差异。因此，真核生物的基因表达可以在 DNA 水平、转录水平、转录后水平、翻译水平以及翻译后水平等多个层次上分别进行调控。

一、真核生物染色质结构直接影响基因转录

以染色质形式组装在细胞核内的 DNA 所携带的遗传信息表达直接受到染色质结构的制约。染色质的转录非活化状态和转录活化状态在结构和性质上有很大的差别，DNA 甲基化和组蛋白共价修饰所导致的染色质的压缩状态造成了 DNA 的不可接近性，它们将直接影响到基因表达的频率或程度。

1. 转录活化的染色质对核酸酶极为敏感　当染色质活化后，常出现一些对核酸酶（如 DNase I）高度敏感的位点，称之超敏位点（hypersensitive site）。超敏位点通常位于被活化基因的 5′侧翼区 1 000 bp 内，但有时也会在更远的 5′侧翼区或 3′侧翼区出现一些超敏位点。这些的转录活化区域是缺乏或没有核小体结合的“裸露”DNA 链。

2. 转录活化的染色质的组蛋白发生改变　转录活跃区域的染色质中的组蛋白的特点

是：①富含赖氨酸的 H1 组蛋白含量降低；②H2A-H2B 组蛋白二聚体的不稳定性增加，使它们容易从核小体核心中被置换出来；③核心组蛋白可以被甲基化、磷酸化、乙酰化以及泛素化修饰。这些都使得核小体的结构变得松弛而不稳定。

负责组蛋白共价修饰的一些蛋白分子，如组蛋白乙酰化酶（histone acetyltransferase，HAT）和组蛋白去乙酰化酶（histone deacetylase，HDAC）在 DNA 水平的基因表达调控中具有重要作用。HAT 使组蛋白发生乙酰化，促使染色质结构松弛，有利于基因的转录，被称为转录辅激活因子（co-activator），而 HDAC 促进组蛋白的去乙酰化，抑制基因的转录，被称为转录辅抑制因子（co-repressor）。

3. RNA 聚合酶结合位点的上游和下游具有不同的超螺旋构象　转录启动后，RNA 聚合酶沿 DNA 向下游移动，使得下游的 DNA 呈现出正超螺旋结构，有利于核小体结构的解体；而 RNA 聚合酶上游的 DNA 结构则以负超螺旋为主，有利于核小体结构的再形成。这样的超螺旋结构差异有助于 RNA 聚合酶向下游移动，进行转录。

4. CpG 岛甲基化水平降低　DNA 甲基化（DNA methylation）是真核生物在染色质水平控制基因转录的重要机制。真核细胞基因组中约有 5%的胞嘧啶在 DNA 甲基转移酶（DNA methyltransferase）的作用下被甲基化为 5-甲基胞嘧啶，并且以序列 CG（称为 CpG 岛）中的胞嘧啶甲基化更加常见。业已发现处于转录活跃状态的染色质中，CpG 岛的甲基化程度下降，例如管家基因的 CpG 岛中胞嘧啶甲基化水平较低。CpG 岛的高甲基化促进染色质形成致密结构，因而不利于基因表达。

染色质结构对基因表达的影响可以遗传给子代细胞，其机制是细胞内存在着具有维持甲基化作用的 DNA 甲基转移酶，可以在 DNA 复制后，依照亲本 DNA 链的甲基化位置催化子链 DNA 在相同位置上发生甲基化。这种现象称为表观遗传（epigenetic inheritance）。这种遗传信息不是蕴藏在 DNA 序列中，而是通过对染色质结构的影响及基因表达变化而实现的。

二、转录速度决定 RNA 合成效率

与原核细胞一样，转录起始是真核生物基因表达调控的关键。但是，真核生物的基因表达起始要比原核细胞的复杂得多，这是因为真核生物的 RNA 聚合酶需要与多个转录因子相互作用而形成转录起始复合物。

（一）顺式作用元件可直接影响基因表达活性

真核生物的顺式作用元件按其功能的不同，可分为启动子、增强子和沉默子。

1. 真核生物启动子结构和调节较原核生物复杂　真核生物启动子的序列一致性不像原核生物那样明显，而且 RNA 聚合酶与 DNA 的结合需要多种蛋白质因子的相互协调作用。因此，真核生物的启动子序列要比原核生物的复杂得多、序列也更长。真核生物启动子一般包括转录起始点及其上游约 100～200 bp 序列，包含有若干具有独立功能的 DNA 序列元件，每个元件约长 7～30 bp。真核生物有三种 RNA 聚合酶，它们分别结合在三类不同的启动子上负责转录不同的 RNA（见第一章和第五章）。

2. 增强子是一种能够提高转录效率的顺式调控元件　它的长度大约是 200 bp，可使旁侧的基因转录效率提高 100 倍。增强子由若干组件构成，基本核心组件常为 8～12 bp，可以单拷贝或多拷贝串联的形式存在。增强子的功能及其作用特征如下：

（1）增强子与被调控基因位于同一条 DNA 链上，属于顺式作用元件。

（2）增强子是组织特异性转录因子的结合部位。因此增强子一般具有组织或细胞特异性，即仅在某些细胞或组织中表现活性。

（3）增强子不仅能够在基因的上游或下游起作用，而且还可以远距离实施调节作用（通常情况为 1～4 kbp），个别情况下甚至可以调控 30 kbp 以外的基因。

（4）增强子作用与序列的方向性无关。将增强子的方向倒置后依然能起作用，而方向倒置后的启动子就不能起作用。

（5）增强子需要有启动子才能发挥作用，没有启动子存在，增强子不能表现活性。但增强子对启动子没有严格的专一性，同一增强子可以影响不同类型启动子的转录。

3. 沉默子能够抑制基因的转录　沉默子是一类负性调控元件，最初在酵母中发现。已有的证据显示沉默子与增强子类似，其作用亦不受序列方向的影响，也能远距离发挥作用，并可对异源基因的表达起作用。

（二）转录因子是转录调控的关键分子

转录因子是在转录过程中所必需的除 RNA 聚合酶以外的蛋白质因子，与反式作用因子含义相同，参加转录起始和延长。真核生物的三种 RNA 聚合酶所需的转录因子有所不同。转录因子一般含有三个不同的功能结构域：DNA 结合结构域（DNA binding domain）、转录激活结构域（activation domain）和与其他蛋白质结合的结构域（protein binding domain）。

1. 转录因子利用特定的蛋白质模体与 DNA 结合

（1）锌指（zinc finger）模体结构：这是一类含锌离子的形似手指的蛋白模体（motif），每个重复的“指”状结构约含 23 个氨基酸残基，形成 1 个 α-螺旋和 2 个反向平行的 β-折叠的二级结构。每个 β-折叠上有 1 个半胱氨酸（Cys）残基，而 α-螺旋上有 2 个组氨酸（His）或半胱氨酸（Cys）残基。这 4 个氨基酸残基与二价锌离子之间形成配位键（图 6-7）。整个蛋白质分子可有多个这样的锌指重复单位。每一个单位可将其指部伸入 DNA 双螺旋的大沟内，接触 5 个核苷酸。例如与 GC 盒结合的人成纤维细胞转录因子 SP1 中就有 3 个锌指重复结构。

（2）碱性螺旋-环-螺旋（basic helix-loop-helix，bHLH）模体结构：这类结构至少有两个 α-螺旋，由一个短肽段形成的环所连接，其中一个 α-螺旋的 N-末端富含碱性氨基酸

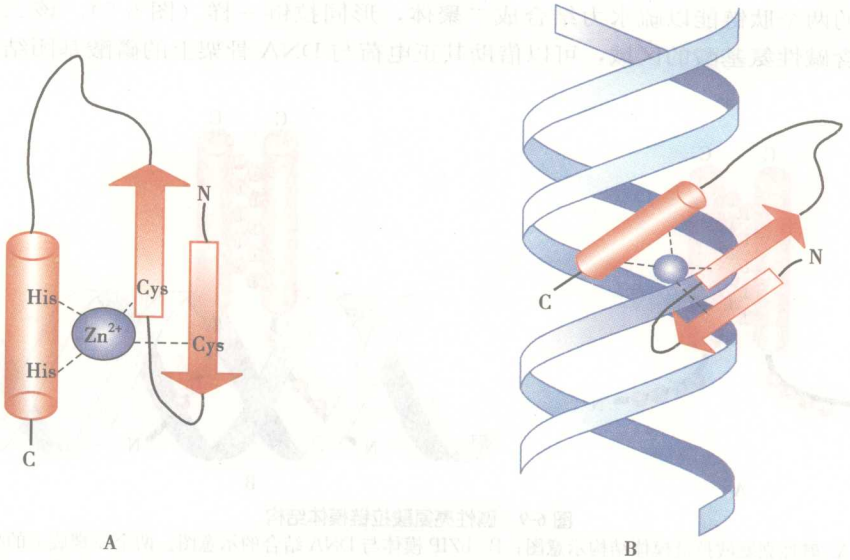


图 6-7　锌指模体结构

A. 独立的锌指模体示意图；B. 锌指模体与 DNA 结合的示意图。锌指模体的 α-螺旋结构结合在 DNA 的大沟中

残基，是与 DNA 结合的结合域（图 6-8）。bHLH 模体通常以二聚体形式存在，而且两个 α-螺旋的碱性区之间的距离大约与 DNA 双螺旋的一个螺距相近（3.4 nm），使两个 α-螺旋的碱性区刚好分别嵌入 DNA 双螺旋的大沟内。

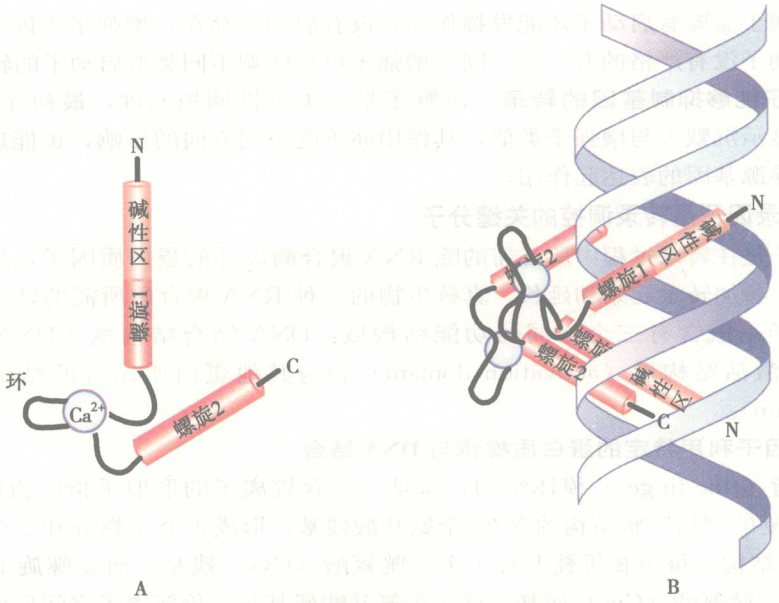


图 6-8　碱性螺旋-环-螺旋模体结构

A. 独立的碱性螺旋-环-螺旋模体结构示意图；B. bLHL 模体二聚体与 DNA 结合的示意图。两个 α-螺旋的碱性区分别嵌入 DNA 双螺旋的大沟内

（3）碱性亮氨酸拉链（basic leucine zipper，bZIP）模体结构：该结构的特点是在蛋白质 C-末端的氨基酸序列中，每隔 6 个氨基酸是一个疏水性的亮氨酸残基。当 C-末端形成 α-螺旋结构时，肽链每旋转两周就出现一个亮氨酸残基，并且都出现在 α-螺旋的同一侧。这样的两个肽链能以疏水力结合成二聚体，形同拉链一样（图 6-9）。该二聚体的 N-末端是富含碱性氨基酸的区域，可以借助其正电荷与 DNA 骨架上的磷酸基团结合。

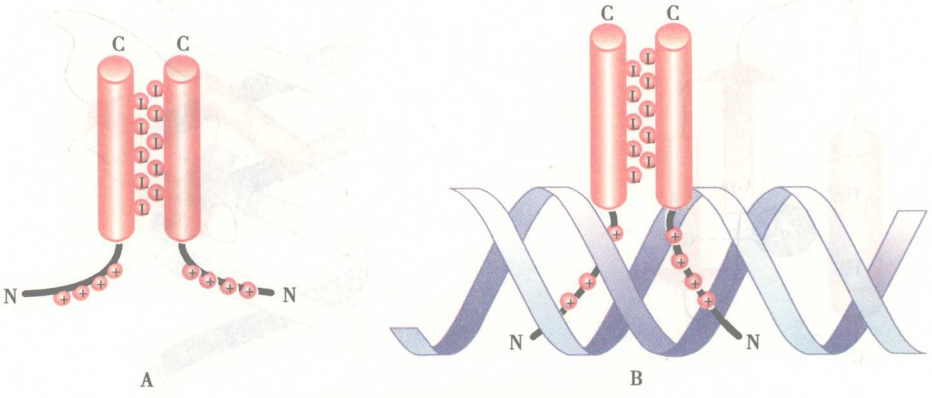


图 6-9　碱性亮氨酸拉链模体结构

A. 碱性亮氨酸拉链模体结构示意图；B. bZIP 模体与 DNA 结合的示意图。两个 α-螺旋上的亮氨酸残基彼此接近，形成了类似拉链的结构，而富含碱性氨基酸残基的区域与 DNA 骨架上的磷酸基团结合

2. 转录激活结构域可增强基因的转录活性　不同的转录因子具有不同的转录激活结构域，根据氨基酸的组成特点，转录激活结构域可分为三类：

（1）酸性激活结构域（acidic activation domain）是一段富含酸性氨基酸的保守序列，常形成带负电荷的β-折叠，通过与 TFⅡD 的相互作用协助转录起始复合物的组装，促进转录。如酵母转录因子 GAL4 的转录激活域。

（2）谷氨酰胺富含结构域（glutamine-rich domain）的 N-末端的谷氨酰胺残基含量可高达 25%左右，通过与 GC 盒结合发挥转录激活作用。

（3）脯氨酸富含结构域（proline-rich domain）的 C-末端的脯氨酸残基含量可高达 20%～30%，通过与 CAAT 盒结合来激活转录。

3. 转录因子在功能上有基本转录因子和组织特异转录因子之分　有些转录因子为 RNA 聚合酶结合启动子所必需，是转录起始所需要的最基本的蛋白组分，被称为通用转录因子（general transcription factors），或称为基本转录因子。真核生物中 3 种不同的 RNA 聚合酶需要的通用转录因子有所不同。催化 mRNA 转录的 RNA 聚合酶Ⅱ有其专用的基本转录因子 TFⅡ A、TFⅡ B 等。

通用转录因子的存在没有组织特异性，因而对于基因表达的时空选择性并不重要。不同的组织细胞则都有其在特定时间、空间所特有的一些转录因子存在，调节特定条件下的基因表达，这些转录因子被称为组织特异转录因子（special transcription factors）。这些组织特异性的转录因子才真正决定着细胞基因的时间、空间特异性表达。图 6-10 列出了一些转录因子的类别及其调节特点。特异性转录因子自身的含量、活性和细胞内定位随时都受到细胞所处环境的影响，是使环境变化在基因表达水平得到体现的关键分子。特异转录因子有的属于转录激活因子（transcription activators），有的属于转录抑制因子（transcription inhibitors）。转录激活因子通常结合在增强子，转录抑制因子通常结合在沉默子。

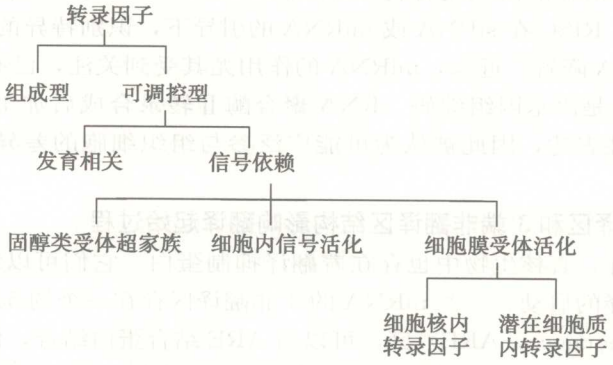


图 6-10　基于功能的转录因子分类

组织特异性转录因子在细胞分化和组织发育过程中具有重要作用。例如，胚胎干细胞的分化方向在相当大的程度上是由细胞内转录因子的种类所决定。阐明各种组织细胞所特有的转录因子种类，就有可能控制细胞的分化方向（见第九章）。

转录调控的实质在于蛋白质-DNA 以及蛋白质-蛋白质之间的相互作用。但是对于生物大分子之间的相互作用及其在生命活动中的具体机制，人们的认识还很肤浅。

三、真核生物在转录后仍可控制 mRNA 的结构和功能

真核生物的基因表达调控在转录后层次不同于原核生物。这一方面是由于两者的转录

产物的剪接、修饰等成熟加工过程有很大的差异，另一方面是由于真核生物的 RNA 产物要被运送至细胞质中去执行功能，其稳定性如何以及其降解过程都可以影响基因表达的最终结果。

（一）mRNA 的稳定性影响真核生物基因表达

mRNA 是蛋白质生物合成的模板，因此它的稳定性将直接影响到基因表达最终产物的数量，是转录后对基因表达进行调控的一个重要因素。真核生物 mRNA 分子的半衰期差别很大，有的可长达数十小时以上，而有的则只有几十分钟或更短。一般而言，半衰期短的 mRNA 多编码调节蛋白，因此，这些蛋白的水平可以随着环境的变化而迅速变化，达到调控其他基因表达的目的。影响细胞内 mRNA 稳定性的因素很多，主要有下面几点：

1. 5′末端的帽子结构可以增加 mRNA 的稳定性　该结构可以使得 mRNA 免于在 5′-核酸外切酶的作用下被降解，从而延长了 mRNA 的半衰期。此外，帽子结构还可以通过与相应的帽子结合蛋白结合而提高翻译的效率，并参与 mRNA 从细胞核向细胞质的转运。

2. 3′末端的 poly（A）尾结构防止 mRNA 降解　poly（A）及其结合蛋白可以防止 3′-核酸外切酶降解 mRNA，增加 mRNA 的稳定性。如果 3′末端 poly（A）被去除，mRNA 分子将很快降解。此外，3′-poly（A）尾结构还参与了翻译的起始过程。实验证明，mRNA 的胞浆定位信号有些也位于 3′-非翻译序列上。组蛋白 mRNA 没有 3′-poly（A）尾的结构，但它的 3′-末端会形成一种发夹结构，使其免受核酸酶的攻击。

3. 一些非编码小分子 RNA 可引起转录后基因沉默　在高等真核生物中，有一类小 RNA（small RNA）由于互补序列的存在，因而能够与 mRNA 结合，导致 mRNA 的降解或翻译抑制，称为转录后基因沉默（post-transcription gene silencing，PTGS）。目前已知的具有 PTGS 作用的小 RNA 有双链的小干扰 RNA（small interference RNA，siRNA）和单链的微 RNA（microRNA，miRNA）两大类。一般认为，两者都可以与细胞内存在的一种被称为 RNA 诱导的沉默复合物（RNA-induced silencing complex，RISC）的核糖核蛋白形成复合物。RISC 在 siRNA 或 miRNA 的引导下，识别特异的 mRNA，并利用其核酸酶活性使 mRNA 降解。近来，miRNA 的作用尤其受到关注，已有近千种 miRNA 被发现。由于 miRNA 是由基因组编码、RNA 聚合酶Ⅱ转录合成后加工而成，且具有组织特异性和阶段特异性表达，因此被认为可能广泛参与组织细胞的差异性基因表达模式的形成。

（二）5′端非翻译区和 3′端非翻译区结构影响翻译起始过程

与原核生物一样，真核生物中也存在着翻译抑制蛋白。它们可以结合在 mRNA 的 5′非翻译区，抑制翻译的启动。一些 mRNA 的 3′非翻译区存在一个约 50 核苷酸长的 AU-丰富序列（AU-rich sequence，ARE）区，可以与 ARE 结合蛋白结合，促使 poly（A）核酸酶切除 poly（A）尾，使 mRNA 降解。因此含有 ARE 区的 mRNA 通常都不稳定。

（三）mRNA 前体的选择性剪接可以调节真核生物基因表达

真核生物基因所转录出的 mRNA 前体含有交替连接的内含子和外显子。通常状态下，mRNA 前体经过剔除内含子序列后成为一个成熟的 mRNA，并被翻译成为一条相应的多肽链。但是，参与拼接的外显子可以不按照其在基因组内的线性分布次序拼接，内含子也可以不完全被切除，由此产生了选择性剪接（见第五章）。选择性剪接的结果是由同一条 mRNA 前体产生了不同的成熟 mRNA，并由此产生了完全不同的蛋白质。这些蛋白质的功能可以完全不同，显示了基因调控对生物多样性的决定作用。

四、翻译水平的调控决定蛋白质合成速度

真核生物翻译的起始过程也是调控蛋白质生物合成过程的关键点。在翻译起始过程中，有大量的翻译起始因子参与，它们的活性变化与基因表达调控密切相关。对起始因子活性的调控，磷酸化与去磷酸化修饰是一种主要的方式，而不同的因子其活性形式各不相同。

1. 翻译起始因子 eIF-2α 的磷酸化抑制翻译起始　eIF-2 是蛋白质合成过程中的一种起始因子，主要参与起始 Met-tRNA$_i$ 的进位过程，其 α 亚基的活性可因磷酸化（cAMP 依赖性蛋白激酶所催化）而降低，导致蛋白质合成受到抑制。如血红素对珠蛋白合成的调节就是由于血红素能抑制 cAMP 依赖性蛋白激酶的活化，从而防止或减少了 eIF-2 的失活，促进了蛋白质的合成。

2. eIF-4E 及 eIF-4E 结合蛋白的磷酸化激活翻译起始　帽结合蛋白 eIF-4E 与 mRNA 帽结构的结合是翻译起始的限速步骤，磷酸化修饰及与抑制物蛋白的结合均可调节 eIF-4E 的活性。磷酸化的 eIF-4E 与帽结构的结合力是非磷酸化的 eIF-4E 的 4 倍，因而可提高翻译的效率。胰岛素及其他一些生长因子都可增加 eIF-4E 的磷酸化从而加快翻译，促进细胞生长。同时，胰岛素还可以通过激活相应的蛋白激酶而使一些与 eIF-4E 结合的抑制物蛋白磷酸化，磷酸化后的抑制物蛋白会与 eIF-4E 解离，激活 eIF-4E。

第四节　基因表达调控异常与疾病

由于转录因子在真核生物基因表达调控中起着至关重要的作用，其结构或功能的改变往往会引起机体的功能缺陷甚至疾病。例如，转录因子 *TBX5* 基因的突变可以引起 HOH-Dream 综合征，导致先天性拇指异常、心脏的房间隔缺损、室间隔缺损或复合型缺损等；再如含有锌指结构的转录因子 GATA3，其基因突变后的临床表现是血液内甲状旁腺素水平降低，导致神经性耳聋和肾发育不良。

不适当的 DNA 共价修饰的也可引起疾病，例如甲基化修饰。基因中特别是基因调节序列中的 CpG 岛甲基化程度降低，可造成相应的基因表达活性升高。在胚胎期正常发育过程中，原癌基因 *c-MYC* 为适应高表达的要求，其调节序列呈现低甲基化状态。但当胎儿出生后，甲基化位点被高度甲基化，*c-MYC* 的表达终止。如果出生后，该基因仍然处于低甲基化状态，则造成原癌基因的激活。某些原发性肝癌患者细胞中的原癌基因 *c-MYC* 就处于低甲基化状态，其甲基化状态与胚胎期相似。此外，在人结肠癌和肺癌细胞中的原癌基因 *c-H-RAS* 和 *c-K-RAS* 也处于低甲基化状态。

再以遗传性维生素 D$_3$ 抵抗性佝偻病为例。由于维生素 D$_3$ 受体基因上与 DNA 结合的部位发生了点突变，使得维生素 D$_3$ 受体的第一个锌指结构中的 Gly 突变为 Asp，第二个锌指结构中的 Asp 和 Arg 都突变为 Gln，导致受体 DNA 结合区结构和电荷的变化，使得受体与 DNA 的结合能力下降。

（关一夫　李雪松）

第七章　基因组学及相关组学

关键知识点

- 基因组学是研究基因组结构和功能的科学。基因组学研究的内容包括基因的结构、组成、存在方式、表达调控模式、基因的功能及相互作用等，是解码生命的所有信息的一门新的前沿科学。基因组学已经对生物学、遗传学、医药学乃至整个人类社会产生了深远影响。
- 人类基因组计划开始于 20 世纪 80 年代中期，2006 年全部完成。该计划完成了人类基因组的遗传图、物理图、转录图和序列图的制作。为理解所有生命活动的分子机制奠定了基础，对医学的发展模式将有革命性的推动。人类基因组计划完成后的后基因组学研究包括功能基因组学、转录组学、蛋白质组学、比较基因组学、糖组学、代谢组学等领域。
- 功能基因组学是在基因组或系统水平上全面地分析基因组中所有基因功能的学科。探讨的是单一细胞在其生命的一定时刻、一定条件下所表达的基因的种类和数量；或比较不同细胞之间、或同一细胞在不同的条件下基因表达的差异。深入认识基因在发育、分化和病理等状态下的功能变化，认识其表达调控方式及调控机制。
- 蛋白质组学是从整体水平研究细胞内蛋白质的组成、功能及其活动规律的科学，旨在阐明生物体全部蛋白质的表达模式、结构模式和功能模式，其研究内容包括分析不同个体、不同细胞在生理和病理状态下全部蛋白质的成分及数量，确定各种组分所在的空间位置、修饰方法、互相作用机制、生物活性和特定功能等，最终揭示蛋白质功能，是基因组 DNA 序列与基因功能之间的桥梁。

关键词： 基因组学；功能基因组学；蛋白质组学；人类基因组计划

DNA 双螺旋结构的发现及“中心法则”的建立和完善，分子生物学的理论体系已经基本建立，DNA 重组技术使人类获得了改造生命的有效手段和途径。目前，生命科学的研究继续在分析和综合两个层面上深入和拓展。虽然传统的基于相对独立的研究小组或实验室的研究活动仍然是主要的工作方式，但随着生命科学研究内涵的丰富和范围的扩大，多个实验室间的合作研究成为当前的新潮流。大规模的跨单位、跨地区、跨国家的联合研究亦呈现强劲的势头。此外，由于复杂系统理论和非线性科学的发展，生命科学研究正从局部观向整体观转变，从线性单向思维走向复杂性思维，从单一学科为主转变为多学科的交叉融合。随着计算生物学和系统生物学等新兴学科的不断涌现，生命科学大综合和大发展的时期即将或已经来临。在此过程中，“组学”的概念应运而生。

研究细胞、组织或整个生物体内某种分子（DNA、RNA、蛋白质、代谢物或其他分子）的所有组成内容的学科，称为“组学”（-omics）。组学研究包括对基因组及基因产物（转录组和蛋白质组）的系统研究，随后必然要上升到细胞机制、分子机制和系统生物学的水平。

在人类基因组计划（human genome project，HGP）实施的短短几年间，各类组学迅速在生命科学界形成（表 7-1）。最早出现的是与 DNA 相关的基因组学（genomics），人类基因组计划完成以后产生了后基因组学（post genomics），随后又形成了许多与各种生物大分子或小分子相关的组学。诸如：基因组学、功能基因组学（functional genomics）、转录组学（transcriptomics）、蛋白质组学（proteomics）、比较基因组学（comparative genomics）、代谢组学（metabonomics）、糖组学（glycomics）等重要领域。

表 7-1　各种组学的名称及其分类

分类角度	组学名称
按照时间先后不同	基因组学、后基因组学
按照研究对象不同	基因组学、转录组学、蛋白质组学、糖组学、代谢组学
按照专业领域不同	营养基因组学/蛋白质组学、免疫基因组学/蛋白质组学、肿瘤基因组学/蛋白质组学、药物基因组学/蛋白质组学、毒理基因组学/蛋白质组学、环境基因组学/蛋白质组学、生殖基因组学/蛋白质组学、表位基因组学等
按照研究目的不同	结构基因组学/蛋白质组学、功能基因组学/蛋白质组学、比较基因组学/蛋白质组学

第一节　基因组学

一、基因组学的概念

基因组学指对所有基因进行基因作图（包括遗传图谱、物理图谱、转录图谱）、核苷酸序列分析、基因定位、基因结构与功能的关系以及基因间相互作用的一门科学，简言之，就是研究基因组结构和功能的科学。基因组学研究的内容包括基因的结构、组成、存在方式、表达调控模式、基因的功能及相互作用等，是解码生命的所有信息的一门新的前沿科学。基因组学已经对生物学、遗传学、医药学乃至整个人类社会产生了深远影响。

二、人类基因组计划

（一）人类基因组计划的提出

人类基因组计划（HGP）最先由美国提出并启动，随后英国、日本、法国、德国、中国等国家相继加入，是描述人类基因组和其他模式生物体基因组特征，在整体上破译遗传信息，发展基因组新技术，并阐明与此相关的伦理、法律和社会影响的一个国际性研究项目。HGP 与“曼哈顿原子弹计划”、“阿波罗登月计划”一起被誉为 20 世纪自然科学史上“最伟大的三个计划”。

测序人类基因组的可能性于 20 世纪 80 年代中期提出。1985 年 6 月，美国能源部提出 HGP 的初步草案，经讨论后于 1986 年 3 月宣布实施这一草案。1986 年 3 月 7 日，美国科学家 R. Dulbecco 在《Science》杂志上发表的短文中率先提出“测定人类的整个基因组序列”的主张，并引起了世界性的反响，不仅推动了美国，也推动了全世界的 HGP 开展。

（二）人类基因组计划的研究内容和任务

HGP 的研究内容体现为完成基因组的 4 张图，即遗传图、物理图、转录图和序列图。1993 年美国国家人类基因组研究中心修订后的 HGP 具体研究内容包括：人类基因组作图及序列分析；基因的鉴定；基因组研究技术的建立、创新与改进；模式生物（包括大肠杆菌、酵母、果蝇、线虫、小鼠、水稻、拟南芥等）基因组的作图和测序；信息系统的建立，信息的储存、处理及相应的软件开发；与人类基因组相关的伦理学、法学和社会影响与结果的研究；研究人员的培训；技术转让及产业开发；研究计划的外延等。

将基因组采用不同的标志和手段进行分解，使之成为小的结构区域而便于测序，这一过程称为作图。根据采用的方法不同，将绘制出的图谱分为以下四类，即遗传图、物理图、转录图和序列图。

1. 遗传图　遗传图（genetic map）又称连锁图（linkage map），通过计算连锁的遗传

标记之间的重组频率，确定它们的相对距离，即以具有遗传多态性的遗传标记作为“位标”，遗传学距离作为“图距”的基因组图，一般用厘摩（cM，即每次减数分裂的重组频率为1%）来表示。

遗传图的绘制需要应用多态性标志。20世纪80年代中期最早应用的标志是限制性酶切片段长度多态性（restriction fragment length polymorphism，RFLP）。80年代后期发展的短串联重复序列（short tandem repeat，STR），又称微卫星（microsatellite）标志。第三代多态性标志，即单核苷酸多态性（single nucleotide polymorphism，SNP）标志近来被大量使用。

2. 物理图 物理图（physical map）是通过测定遗传标志的排列顺序与位置而绘制成的，即以一段已知核苷酸的DNA片段为“位标”，以DNA实际长度（Mb或kbp）为“图距”的基因图谱。

HGP在整个基因组染色体每隔一定距离标上序列标签位点（sequence tagged site，STS）之后，随机将每条染色体酶切为大小不等的DNA片段，以酵母人工染色体（yeast artificial chromosome，YAC）或细菌人工染色体（bacterial artificial chromosome，BAC）等作为载体构建YAC或BAC邻接克隆系，确定相邻STS间的物理联系，绘制以Mb、kbp、bp为图距的人类全基因组物理图谱。

3. 转录图 转录图（transcription map）又称cDNA图或表达图（expression map），是一种以表达序列标签（expressed sequence tags，EST）为位标绘制的分子遗传图谱。通过从cDNA文库中随机挑取的克隆进行测序所获得的部分cDNA的5′或3′端序列称为表达序列标签，一般长300～500 bp左右。将mRNA反转录合成的cDNA或EST的部分cDNA片段作为探针与基因组DNA进行分子杂交，标记转录基因，绘制出可表达基因转录图，最终绘制出人体所有组织、所有细胞以及所有发育阶段的全基因组转录图谱。EST不仅为基因组遗传图谱的构建提供了大量的分子标记，而且来自不同组织和器官的EST也为基因的功能研究提供了有价值的信息。此外，EST还有助于基因的鉴定，通过分析基因组序列能够获得基因组结构的完整信息，如基因在染色体上的排列顺序，基因间的间隔区结构，启动子的结构以及内含子的分布等。截至2008年3月，美国国立卫生与健康研究院的数据库中收集的人的EST数目是8 136 592个。

4. 序列图 序列图（sequence map）即人类基因组核苷酸序列图，是人类基因组在分子水平上最高层次、最详尽的物理图。其绘制方法是在遗传图谱和物理图谱基础上，精细分析各克隆的物理图谱，将其切割成易于操作的小片段，构建YAC或BAC文库，得到DNA测序模板，测序得到各片段的碱基序列，再根据重叠的核苷酸顺序将已测定序列依次排列，获得人类全基因组的序列图谱。2000年6月美国、英国、法国、德国、日本与中国几乎同时宣布人类基因组计划工作草图的完成，成为生命科学研究的一个里程碑。

上述人类基因组计划的四张图谱被誉为人类“分子水平上的解剖图”（图7-1），也被称为“生命元素周期表”，为本世纪医学和生物学的进一步发展与新的飞跃奠定了基础。

（三）人类基因组计划的完成情况

2000年6月26日，国际人类基因组测序联盟（IHGSC）与Celera公司联合发布了“人类基因组工作草图”；《Nature》和《Science》杂志在2001年2月几乎同时分别公布了人类基因组“精确图”（准确度达到99.99%）及初步分析结果，两组数据存在一定差异，但大部分高度吻合，且发现人类基因数目约为3.4万～3.5万个，仅比果蝇多2万个，远小于原来估计的10万个基因；2003年4月，在DNA双螺旋结构发表50周年之际，人类基因组计划基本完成。

2004年10月21日，国际人类基因组测序联盟在《Nature》杂志上公布了人类基因组

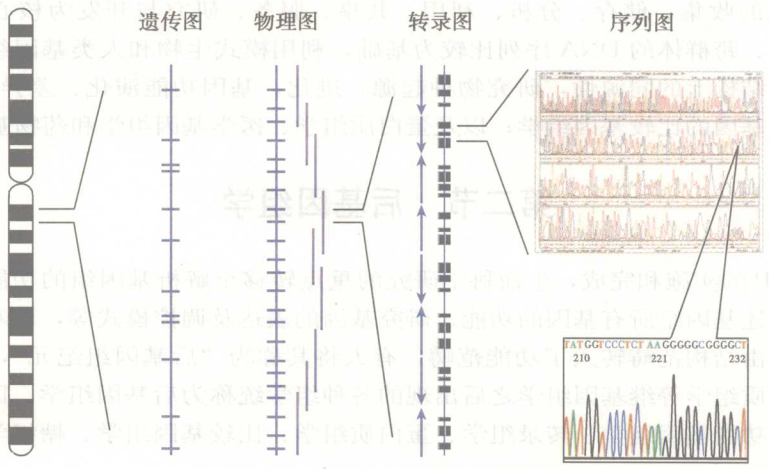


图 7-1　人类基因组计划要完成的四张图谱

的完成图，认为人类只有 2 万～2.5 万个基因（目前认为约 2.8 万个），比 2001 年的估计还要少。

美国和英国科学家在 2006 年 5 月 18 日于《Nature》杂志网络版上发表了人类最后一个染色体——第 1 号染色体的基因测序，标志着解读人体基因密码的“生命之书”宣告全部完成。

作为 HGP 第六个成员国，我国于 1999 年 10 月 1 日正式启动 HGP，2000 年 4 月完成了 1%人类基因组“工作框架图”；2001 年 8 月 26 日绘制完成“中国卷”，提前 2 年获得精确度达 99.99% 的“完成图”序列。

（四）人类基因组计划的意义

人类基因组计划实现了人类基因组的破译和解读，对于认识各种基因的结构和功能，了解基因表达及调控方式，理解生物进化的基础，进而阐明所有生命活动的分子基础具有十分重要的意义。

（1）鉴定人类的全部基因，推动生物技术的进一步发展：HGP 所获得的庞大的 DNA 序列，将对生物技术的研究提供指导性依据。HGP 的大规模运作也将推动生物技术科研与开发并肩走向操作的规模化、自动化，这无疑会使生物技术在未来经济发展中占有非常重要的位置。

（2）将把人类带入基因医学的新时代：随着基因组“工作框架图”的问世，许多致病单基因被确定为候选基因，现已成功分离到亨廷顿舞蹈症、杜氏肌营养不良、哮喘病、乳腺癌等 70 多种遗传病的致病基因。目前，多基因遗传病已成为疾病基因组学研究的重点。遗传性疾病的基因定位，尤其是对多基因复杂性状的基因位点将可以进行全基因组的定位扫描，使得确定致病基因的工作更为容易。例如癌症、高血压、糖尿病等都在吸引着众多的医学家和药物学家从分子水平突破对这些疾病的传统认识，从而改变治疗方式。HGP 使人类在了解致病机制和发现新基因上迈出了至关重要的一步，这为基因药物的设计提供了重要的理论基础和设计原则。

（3）推动了模式生物基因组的研究：人类基因组“工作框架图”的建立带动了小鼠、大肠杆菌、酵母菌、美丽线虫、果蝇等模式生物以及近 50 多种微生物全基因组破译和一些其他生物 DNA 图谱的绘制，也推动了 HGP 向纵深发展，可为人类致病基因的研究提供有价值的参考。

（4）促进了学科交叉与重组：HGP 的深入发展诞生了许多新学科和新领域，其中包

括以生物信息的收集、储存、分析、利用、共享、服务、研究与开发为核心的生物信息学；以跨物种、跨群体的DNA序列比较为基础，利用模式生物和人类基因组织之间编码序列和组成、结构上的同源性，研究物种起源、进化、基因功能演化、差异表达和定位、克隆人类疾病基因的比较基因组学；以及蛋白质组学、医学基因组学和药物基因组学等。

第二节　后基因组学

随着HGP的实施和完成，生命科学研究的重点转移至解析基因组的功能，包括详尽分析序列、描述基因组所有基因的功能、研究基因的表达及调控模式等，即对基因组的研究已经将重心由结构范畴转入了功能范畴，有人将其称为“后基因组纪元”，而将功能基因组学及蛋白质组学等继基因组学之后出现的各种组学统称为后基因组学。因此，后基因组学应该包括功能基因组学、转录组学、蛋白质组学、比较基因组学、糖组学、代谢组学等学科领域。

一、功能基因组学

功能基因组学是利用基因组学提供的信息和产物，发展和应用新的实验手段，在基因组或系统水平上全面地分析基因组中所有基因功能的学科。它以揭示基因组功能及调控机制为目标，探讨的是单一细胞在其生命的一定时刻、一定条件下所表达的基因的种类和数量；或比较不同细胞之间、或同一细胞在不同的条件下基因表达的差异。这一研究将促进对细胞的分化及其功能的理解，同时将了解不同生理病理条件下的基因表达状态，从而深入认识基因在发育、分化和病理等状态下的功能变化，认识其表达调控方式及调控机制。功能基因组学的主要研究包括以下内容：

1. 识别和鉴定人类基因组　识别和鉴定基因序列是提取、鉴定人类基因组必不可少的基础工作。基因的识别和鉴定需采用计算机生物学、生物信息学（bioinformatics）和分子生物学实验手段相结合的方法。同时，还需要利用收集的现有数据，不断扩大数据库，并研制、建立更多样化的数据库和分析软件。

2. 分析基因功能　主要的研究策略是利用计算机进行同源搜索，根据已知序列、进化相关性，发现重要的蛋白质功能域；也可对基因系列进行突变或剔除后，结合功能、表型变化的实验鉴定基因的功能。

3. 研究基因表达谱　基因表达谱包括转录组谱和蛋白质表达谱，反映了一定环境、一定细胞类型、一定细胞生长阶段和一定细胞状态下基因功能的信息。制作的基因表达谱应当包括每种组织、每种细胞的基因表达谱，细胞在不同发育阶段的基因表达谱，正常和疾病状态下的基因表达谱，以及治疗状态下的基因表达谱等。可以利用基因芯片（DNA chip）技术、基因表达系列分析（serial analysis of gene expression，SAGE）、差异显示PCR、消减杂交法等技术分析基因表达情况，从而绘制基因表达谱。

功能基因组学的研究将为人们深入理解人类基因组遗传语言的逻辑构架，基因结构与功能的关系，个体发育、生长、衰老和死亡机制，神经活动和脑功能表现机制，细胞增殖、分化和凋亡机制，信息传递和作用机制，疾病发生、发展的机制和各种生命科学问题提供共同的科学基础。功能基因组研究成果，不仅具有巨大的科学意义，而且有着十分广泛的应用前景。在医疗卫生方面，研究成果可用于医药的发现和开发：致病基因或疾病易感基因的鉴定和克隆，可用于全新原理的诊断、治疗和预防方法的设计；医生将能够根据患者的个体遗传信息，设计个体化的药物疗法；科学家们在人体器官和组织“重造”以及

修复方面将取得巨大进步；以基因组成果为基础的基因组工业，将带动一批高新技术产业向新的领域开拓。

二、转录组学

功能基因组学在描述基因表达模式时涉及一个重要概念——转录组。转录组（transcriptome）是指一种生物基因组表达的全部转录产物的总称，包括某一环境条件、某一生命阶段、某一生理或病理（功能）状态下，生命体的细胞或组织所表达的基因种类和水平。以转录组分析为研究内容的研究领域称为转录组学，即研究细胞在某一功能状态下所含 mRNA 的类型与拷贝数。转录组学的研究方法通常包括基因芯片技术、基因表达系列分析及大规模平行测序系统（massively parallel signature sequencing）。

与基因组不同，转录组的定义中包含了时间和空间的限定。同一细胞在不同的生长时期及生长环境下，其基因表达情况是不完全相同的。而在相同的生长时期及生长环境下，不同组织的转录组也是不同的。

转录组学是功能基因组学的重要分支，也是连接基因组结构和功能一个桥梁和纽带，更是基因调控研究的主要基础和层面。通过系统地研究转录组而得到转录组谱，可以提供生物的哪些基因在何时何种条件下表达或不表达的信息，这些信息能用于推断相应未知基因的功能，或补充已知基因的功能，可以揭示特定调节基因的作用机制，从而有利于更深入地了解基因表达的调控机制。通过这种基于基因表达谱的分子标签，不仅可以辨别细胞的表型归属，还可以用于疾病的诊断。

第三节　蛋白质组学

虽然目前可以利用基因芯片、基因表达系列分析等可以在转录组学层次了解基因表达的基本情况，但大部分生命活动的真正执行者是蛋白质。细胞内蛋白质的种类、数量、化学修饰（如糖基化、磷酸化、泛素化）是细胞行为及调节的重要结构基础。另外，蛋白质与蛋白质的相互作用、蛋白质与其他生物分子的相互作用等活动，也无法在基因组和转录组水平上获知。另一方面，人类以往的视角只是针对生命活动中某一种或某几种蛋白质，难以从整体透彻地阐释生命活动的基本机制。因此，无论是从基因组计划的局限性，还是从蛋白质研究的自身发展而言，大规模、全方位的蛋白质研究均是势在必行。在这一背景下，20 世纪 90 年代中期，国际上萌发了一门在整体水平上研究细胞内蛋白质的组成及其活动规律的新兴学科——蛋白质组学，并作为后基因组学的重要组成部分，系统全面地研究解析蛋白质组。

一、蛋白质组和蛋白质组学的概念

蛋白质组（proteome）一词，源于蛋白质与基因组两个英文单词的组合。广义概念指“一种基因组所表达的全套蛋白质”。蛋白质组学是从整体水平研究细胞内蛋白质的组成、功能及其活动规律的科学，旨在阐明生物体全部蛋白质的表达模式、结构模式和功能模式，其研究内容包括分析不同个体、不同细胞在生理和病理状态下全部蛋白质的成分及数量，确定各种组分所在的空间位置、修饰方法、互相作用机制、生物活性和特定功能等，最终揭示蛋白质功能，是基因组 DNA 序列与基因功能之间的桥梁。

二、蛋白质组与基因组的比较

每一种生物只有一种基因组，但不同组织、细胞却能表达出多样的蛋白质组。在一个特定的细胞内或在不同的细胞之间，蛋白质的存在形式可能不同；即使是同一细胞，在不同时期、不同的生长条件（正常、疾病或外界环境刺激）下表达的蛋白质也不同。要系统地进行蛋白质组学的研究，就应该认识到蛋白质组的特点及它与基因组存在的重要区别（表 7-2）。

表 7-2 基因组与蛋白质组的比较

	基因组		蛋白质组	
研究对象	DNA		蛋白质	
稳定性	相对稳定，静态		变化大，动态	
研究内容及方法	结构基因组	序列分析	结构蛋白质组	晶体结构分析
		多态性作图		生物质谱
	功能基因组	表达系列分析	功能蛋白质组	双向电泳、生物质谱
		基因芯片		蛋白芯片
		转基因技术		酵母双杂交技术

1. 蛋白质组具有多样性 在转录时，一个基因可有多种 mRNA 选择性剪接，而且同一蛋白又可能以多种形式进行翻译后的修饰，故一个蛋白质组不是一个基因组的直接产物。迄今为止发现的蛋白质翻译后修饰方式已超过 200 种，常见的有磷酸化、糖基化、酰基化、硝基化、磺基化、脂化、泛素化和水解修饰等。如果把一种修饰蛋白视为一种新的蛋白质，那么蛋白质组的蛋白质数量将远远大于相应的基因组的基因数量。如果将修饰模式考虑在其中，目前估计人类蛋白质组的蛋白质种类大约在 20 万到 200 万之间。

从蛋白质修饰的角度来看，不仅仅是蛋白质种类大大增加，更重要的是，由于度量蛋白质修饰种类的尺度尚不存在，人们也许永远不能像确定基因组核苷酸序列那样，准确地统计出生物体内蛋白质组的蛋白质总数。从这种意义上来说，对基因组核苷酸序列的测定是一种“有限”的工作，而对蛋白质组蛋白质种类的确定则是一种相对“无限”的工作。

2. 蛋白质组随时间和空间而动态变化 在个体发育的不同阶段或细胞的不同活动时期，细胞内产生的蛋白质种类是不一样的。不同蛋白质的寿命也不一样。有些蛋白质在合成后成为细胞的结构成分，因此相当稳定；而有些蛋白质在产生后被用来进行某种细胞活动，比如基因转录的调控，工作一旦完成就被迅速降解。因此，在分析蛋白质组的蛋白质成分时，需要把时间作为一个重要的参数。蛋白质组的另一个重要特征是，不同的蛋白质通常分布在细胞的不同部位，它们的功能与其空间定位密切相关。要想真正了解蛋白质的功能，通常还需要知道蛋白质所处的空间位置。更为重要的是，许多蛋白质在细胞中处在一个动态变化的过程，它们常常通过在不同亚细胞环境里的运动发挥作用。例如细胞周期的调控、细胞的信号转导和转录调控等过程，都依赖于蛋白质空间位置的变化和运动。

3. 蛋白质彼此间有着直接的影响 某一个蛋白质功能的实现，通常离不开它与其他蛋白质之间的相互作用。换言之，不与其他蛋白质发生作用的“孤立蛋白质”根本就不存在。蛋白质之间的相互作用大致有三类：与生命活动相关的蛋白质相互作用网络；结构型或功能型蛋白质复合体的形成，包括多亚基蛋白质、多成分的蛋白质复合物等；控制着重要细胞内活动的、瞬时的蛋白质相互作用。

4. 蛋白质组学研究技术尚不成熟　蛋白质组研究需要把时间和空间的要求作为重要的参数来建立相应的技术平台和确定研究工作的技术策略，对技术的依赖性和要求更高。为了解决蛋白质组学研究中的分离和检测问题，到目前为止，虽然人们已经从化学、生物化学、分子生物学、免疫学、结构生物学等多个层面提出了许多技术策略，但是还没有建立起像“人类基因组测序”那样的高效可行的技术平台。

三、蛋白质组学研究的内容和方法

由于蛋白质组本身结构和功能的复杂性，研究手段的不完善等原因，使得蛋白质组学的发展远远落后于基因组学和转录组学。目前可利用的蛋白质组学技术包括双向凝胶电泳和新的生物质谱技术等。

蛋白质组学的研究主要包括两个方面的内容：一是对蛋白质表达模式的研究，即对蛋白质组成的研究，又称为结构蛋白质组学（structural proteomics），可利用双向凝胶电泳、质谱、高性能的X-射线晶体衍射及核磁共振等技术，研究蛋白质的结构（包括修饰）、定位、移位以及对蛋白质组组分进行分析鉴定。其中，双向凝胶电泳和质谱技术是当前分离鉴定蛋白质的两大支柱技术；二是对蛋白质功能模式的研究，也称为功能蛋白质组学（functional proteomics），可通过系统地利用中和抗体、小分子化合物等方法干预蛋白质的活性或使其失活，观察对某一生命现象的影响，从而直接描述该蛋白质的功能，或利用酵母双杂交、噬菌体展示技术等研究蛋白质相互作用，有人也将此称为相互作用蛋白质组学（interactional proteomics）。目前对蛋白质组功能模式的研究主要还是集中于蛋白质相互作用网络关系，常用酵母双杂交系统。

总体看来，在所有的蛋白质组学相关技术中，双向凝胶电泳技术、质谱技术和计算机图像分析与大规模数据处理技术仍然是蛋白质组研究的三大基本支柱技术。

四、蛋白质组数据库

由于蛋白质组学具有高效率、高通量的特点，它与生物信息技术是密切相关的。蛋白质组数据库是蛋白质组研究水平的标志和基础。目前已有众多与蛋白质组研究相关的数据库（表7-3），其中应用最多的包括蛋白质序列数据库（SWISS-PROT，TrEMBL）、基因序列数据库（GenBank，EMBL）、蛋白模式数据库（Prosite）、蛋白二维凝胶电泳数据库、蛋白三维结构数据库（PDB，FSSP）、蛋白翻译后修饰数据库（O-GLAYCBASE）、基因组数据库（GDB，OMIM）和代谢数据库（ENZYME）等。随着基因组学的迅速推进，会给蛋白质组研究提供更多更全的数据库。

表7-3　常用蛋白质数据库

Swiss-prot/TrEMBL	http：//www. expasy. ch/
GenBank	http：//www. ncbi. nih. gov/Genbank/
EMBL	http：//www. ebi. ac. uk/embl. html
Prosite	http：//www. expasy. ch/sprot/prosite. html
PDB	http：//www. pdb. bnl. gov/
FSSP	http：//www. embl-ebi. ac. uk/dali
O-GLYCBASE	http：//www. cbs. dtu. dk/databases/OGLYCBASE
GDB	http：//www. gdb. org
OMIM	http：//www. ncbi. nih. gov/OMIM
ENZYME	http：//www. expasy. ch/sprot/enzyme. html

五、蛋白质组学的应用及进展

目前，蛋白质组学研究成果主要可应用于以下几个方面：①用于寻找疾病相关的蛋白质，如疾病蛋白质标志物；②用于微生物蛋白组研究，彻底阐明病原微生物的致病机制并寻找全新的药物作用靶点；③蛋白质组数据库将成为药物设计的路标；④对不同生物的蛋白质组进行比较性研究，可以对生物进化途径提供参考，对多细胞生物的起源提供线索；⑤追踪胞内信号分子的移位，阐明目标蛋白质在信号转导途径中的位置。

在基础研究方面，近两年来蛋白质组研究技术已被应用到各种生命科学领域，如细胞生物学、神经生物学等。在研究对象上，覆盖了原核微生物、真核微生物、植物和动物等范围，涉及各种重要的生物学现象，如信号转导、细胞分化、蛋白质折叠等。在未来的发展中，蛋白质组学的研究领域将更加广泛。

在应用研究方面，蛋白质组学将成为寻找疾病分子标记和药物靶标最有效的方法之一。在对癌症、早老性痴呆等人类重大疾病的临床诊断和治疗方面蛋白质组技术也有十分诱人的前景，目前国际上许多大型药物公司正投入大量的人力和物力进行蛋白质组学的应用性研究。

在技术发展方面，蛋白质组学的研究方法将出现多种技术并存，各有优势和局限的特点。除了发展新方法外，更强调各种方法间的整合和互补，以适应不同蛋白质的不同特征。另外，蛋白质组学与其他学科的交叉也将日益显著和重要，特别是蛋白质组学与其他大规模科学，如基因组学、生物信息学等领域的交叉，所呈现出的系统生物学（system biology）研究模式，将成为未来生命科学最令人激动的新前沿。

第四节　后基因组时代生命科学的发展趋势

后基因组学时代生物医学研究模式将呈现八大变化趋势，即：①由基因组学向功能基因组学转变；②由基因组学向蛋白质组学转变；③由以作图为基础的基因鉴定向以序列为基础的基因鉴定转变；④由单基因病研究向多基因病研究转变；⑤由对疾病的特异性DNA诊断向疾病易感性监测转变；⑥由分析单个基因向分析基因家族、生化通路和系统中多基因转变；⑦由研究基因的作用向研究基因作用的调控机制转变；⑧由研究疾病的病因（特异性突变）向研究疾病的病理发生（机制）转变。

一、基因组学研究的深入

2005年，国际人类基因组单体型图（HapMap）计划的顺利完成是基因组学研究的又一突破。HapMap计划旨在确定和编目人类遗传的相似性和差异性。利用HapMap计划获得的信息，研究人员将能够发现与人类健康、疾病以及对药物和环境因子的个体反应差异相关的基因。来自美国、中国、英国、日本、加拿大等国的科学家于2005年10月成功地完成了人类基因组单体型图的绘制。在“HapMap”工程中，由中科院北京基因组所、国家人类基因组南方研究中心、香港大学的科学家携手绘制了“HapMap”计划10%的“中国卷”部分，构建了第3号、第21号染色体和第8号染色体短臂的单体型图，为进一步研究我国各人群的基因组多态性及其对疾病易感性和药物反应差异性的影响打下了坚实的基础。

在 HapMap 基础上，美国国立癌症研究所与国立基因组研究所于 2005 年共同公布并启动了人类癌症基因组计划，该计划旨在找到所有致癌基因的微小变异，绘制癌症基因图谱，为癌症的诊断、预防与治疗提供线索。该计划将对 1.25 万份肿瘤样本进行基因测序，涉及 50 种类型癌症，测序工作规模预计比人类基因组计划大得多。

二、基因组学相关领域

随着基因组学研究的不断深入，人类有望揭示生命物质世界的各种前所未知的规律，完全揭开生命之谜，进而驾驭生命，使之为人类的社会经济服务。基因组研究和其他学科研究交叉，促进一些新的学科领域的诞生，如生物信息学、比较基因组学、代谢组学、糖组学、免疫组学（immunomics）、营养基因组学（nutritional genomics）、环境基因组学（environmental genomics）、药物基因组学（pharmacogenomics）、病理基因组学（pathogenomics）、生殖基因组学（reproductive genomics）、群体基因组学（population genomics）等。

1. 生物信息学　生物信息学是以生物大分子为研究对象，以计算机为工具，运用数学和信息科学的观点、理论和方法去研究生命现象，组织和分析呈指数级增长的生物信息数据的一门科学。其研究的重点体现在基因组和蛋白质组两个方面，关注的研究热点包括：序列对比，基因识别和 DNA 序列分析，蛋白质结构预测，分子进化，数据库中知识发现（knowledge discovery in database）。生物信息学是基于数据库与知识发现的研究，当前已成为基因组科学中的带头学科，对生命科学带来革命性的变化。

2. 比较基因组学　比较基因组学是基于基因组图谱和测序基础上，对已知的基因和基因组结构进行比较，来了解基因的功能、表达机制和物种进化的学科。与功能基因组学一样，比较基因组学主要聚焦在基因组水平，分析两个或两个以上的物种，目的是发现（不同物种基因序列）或非基因序列间广泛、特异的相似性和差异，揭示不同物种基因组的进化关系。由于直接人体实验的伦理和法律问题，所以比较基因组学为通过研究模式生物基因的功能来确定人类基因的功能提供了有价值的手段。比较基因组学可利用模式生物基因组与人类基因组之间编码顺序上和结构上的同源性，克隆人类疾病基因，揭示基因功能和疾病分子机制，阐明物种进化关系及基因组的内在结构。

3. 代谢组学　代谢组学通过考察生物体系受刺激或扰动后（如将某个特定的基因变异或环境变化后）其代谢产物的变化或其随时间的变化，研究生物体系的代谢途径。与基因组学、转录组学和蛋白质组学相对应，代谢组学是一门对某一生物或细胞所有低相对分子质量代谢产物进行定性和定量分析，以监测活细胞中化学变化的学科。代谢组学研究的重要性主要表现在：①药物作用（药效和毒性）模型的鉴别和确证；②药物作用机制的研究；③药物的临床前毒性评价；④药物的临床前安全性评价；⑤疾病诊断。

4. 糖组学　20 世纪末，继基因组学、蛋白质组学之后，糖组学也日益受到人们关注。糖组学研究糖链组成及其功能，是基因组学的后续和延伸。具体内容包括研究糖与糖之间、糖与蛋白质之间、糖与核酸之间的联系和相互作用。糖组学主要针对糖蛋白，涉及单个个体的全部糖蛋白结构分析，确定编码糖蛋白的基因和蛋白质糖基化的机制。

糖组学是后基因组学的一个重要领域，许多生物信息都体现在丰富多样的复合糖形式，综合应用基因组、蛋白质组、代谢物组、糖组的概念来全面解释复杂的生命系统是完全必要的，随着糖蛋白分离技术和相关仪器的发展，糖组学将进一步丰富和发展。

5. 免疫组学　免疫组学和免疫组研究免疫反应相关的全套分子、它们的作用靶分子及其功能。免疫组学包括了免疫基因组学（immunogenomics）、免疫蛋白质组学（immu-

noproteomics）和免疫信息学（immunoinformatics）三方面的内容，特别强调在基因组学和蛋白质组学研究的基础上，充分利用生物信息学、生物芯片、系统生物学、结构生物学、高通量筛选等技术，大规模开展免疫系统和免疫应答分子机制研究，发现新免疫相关分子的功能，为系统全面地了解免疫系统和免疫应答提供基础。

三、基因组与后基因组研究将引发医学革命

由于基因组与后基因组研究的发展，基因医学的概念应运而生。医学家将充分应用基因组与后基因组研究所提供的各种信息和知识储备，解决人类重大疾病的诊断、治疗和预防中的各类问题。

1. 重要疾病机制的阐明　人类常见疾病中，恶性肿瘤、心血管疾病、神经系统退行性疾病、自体免疫病及代谢性疾病等均涉及基因的先天性缺陷与后天的基因突变，以及细胞生长、分化与凋亡之间的调控失衡。随着人类所有基因的结构与功能的阐明，以及对维持人体基因组稳定性、基因表达调控、DNA 复制与修复的物质基础的了解，人类对疾病发生的机制将有重要发现，甚至在概念上有重要更改，从而在预防、诊断、治疗上有新的突破。

2. 基因诊断技术的发展　医学上对许多病原体已有可应用基因诊断技术，未来将出现更为精细的基因诊断方法。病原体诊断将发展为病毒株检测以及致病力强度的检测和抗菌株的发现。多基因疾病（肿瘤、糖尿病、心脑血管疾病等）将采用多基因检测、易感基因检测、代谢酶基因检测、DNA 修复基因检测等，对于预防或早诊早治有重要价值。

3. 基因治疗的推广应用　科学家们已经开始尝试对越来越多的患者进行基因治疗，已涉及肿瘤、艾滋病、肝炎和某些单基因疾病，治疗方案也逐步完善。然而，目前的方案多尚在试验阶段，仍无法满足临床治疗的需要。除了基因表达与转染体系的功效不足外，目的基因的选择仍是主要的难题。随着基因医学的发展，功能基因与疾病相关基因的进一步阐明，有针对性的个体化基因治疗将成为临床的一个常规治疗手段。

4. 基因药业的蓬勃发展　1973 年重组 DNA 技术建立和 1982 年第一个基因工程药物上市以来，全世界范围内已上市基因药物百余种。基因序列结构和功能的阐明将使基因药物的种类大幅增加，新的细胞因子、生长因子及可溶性受体、生长抑制因子、细胞凋亡信号分子、多肽激素等可溶性生物活性物质，可以作为药物直接使用；各种细胞膜受体及胞内调控因子，可作为靶分子设计小分子或其他多肽药物或反义药物；转基因技术使体细胞治疗、器官移植和动物药厂均成为可能。

（宋方洲　秦　琴）

第八章　细胞信号转导

关键知识点

- 细胞通讯和细胞信号转导是机体内一部分细胞发出信号，另一部分细胞接收信号并将其转变为细胞功能变化的过程。
- 细胞信号转导的相关分子包括细胞外信号分子、受体、细胞内信号转导分子。
- 受体的基本类型包括细胞内受体和膜表面受体两大类。膜受体又进一步根据结构与信号转导方式的不同有离子通道型、G蛋白偶联型和蛋白激酶偶联受体三个亚类。受体的功能是结合配体并将信号导入细胞。
- 各种信号转导分子的特定组合及有序的相互作用，构成了不同的信号转导途径。信号转导分子通过引起下游分子的数量、分布或活性状态变化而传递信号。小分子信使以浓度和分布的迅速变化为主、蛋白质信号转导分子依赖蛋白质的相互作用为主而传递信号。
- 信号的传递和终止、信号转导过程中的级联放大效应、信号转导途径的通用性和特异性、信号转导途径的交互联系形成了细胞信号转导的基本规律。

关键词： 细胞信号转导；受体；G蛋白；第二信使；钙调蛋白；蛋白激酶偶联受体；MAPK；SH2结构域；PI3K

生物体内各种细胞在功能上的协调统一是通过细胞间相互识别和相互作用的机制来实现的，这种机制称为细胞通讯（cell communication）。各种细胞或者识别与之相接触的细胞，或者识别所处微环境中的各种化学和物理信号，并将其转变为细胞内各种分子活性的变化，从而能够根据外来信号改变细胞内的某些代谢过程，或改变生长速度，甚至进入程序化死亡过程（凋亡）。外源信息传入细胞内并引起细胞应答反应的过程称为信号转导。细胞通讯和细胞信号转导是机体内一部分细胞发出信号，另一部分细胞接收信号并将其转变为细胞功能变化的过程，这种生命活动过程使生物体内各种细胞在功能上协调统一，使机体在整体上对外界环境的变化发生最为适宜的反应。

第一节　细胞信号转导的分子机制

信号转导（signal transduction）是通过多种分子相互作用的一系列有序反应、将来自细胞外的信息传递到细胞内各种效应分子的过程。各种信号转导分子的有序组合及相互作用构成了不同的信号转导途径（signal transduction pathway）。在信号转导过程中，不同的分子以特定的方式与其所处信号转导途径中的上游分子和下游分子相互作用。

一、细胞信号转导需要多种分子的相互作用

（一）细胞外信号分子在细胞间传递信号

大部分细胞是通过膜表面分子和化学信号分子这两大类细胞外信号分子引起功能的变化或调整。

1. 膜表面信号分子需要细胞间接触才能传递信号　每个细胞都有众多的蛋白、糖蛋

白、糖脂分子分布于质膜的外表面。对与其他靶细胞而言，这些表面分子可作为细胞外信号分子（配体），与相邻靶细胞膜表面分子（受体）特异性地结合，将信号传入靶细胞内，促使靶细胞发生各种功能变化。

2. 化学信号分子作为游离分子在细胞间传递信号　一些细胞可以分泌化学物质（如蛋白质或小分子有机化合物）至细胞外，这些分子作用于靶细胞表面或细胞内的受体，调节靶细胞的功能，因而称为化学信号分子。化学信号分子可根据其物理性质分为脂溶性化学信号和水溶性化学信号两大类。从其作用距离考虑则可分为激素、神经递质和属于旁分泌系统的细胞因子、生长因子等三大类。有些细胞因子分泌后不仅作用于其他细胞，还可以作用于分泌这些因子的细胞本身，因此称为自分泌。

（二）受体结合细胞外信号分子并向细胞内传递信号

细胞外的信号分子都是通过位于靶细胞表面或内部的特殊蛋白质——受体（receptor）将信号导入靶细胞内。

1. 各类受体具有不同的结构和作用特点　受体依据空间定位分为两大类，即细胞内受体和膜表面受体。细胞内受体包括位于细胞质或胞核内的受体，其相应配体是脂溶性信号分子，如类固醇激素、甲状腺素、维甲酸等。细胞外水溶性信号分子和细胞表面信号分子（如生长因子、细胞因子、水溶性激素分子、黏附分子等）不能进入靶细胞，其受体位于靶细胞的细胞质膜外表面。根据结构、接收信号的种类、转换信号的方式等，膜表面受体分为配体门控受体型离子通道（ligand-gated receptor channel）、G 蛋白偶联受体（G protein-coupled receptor，GPCR）和蛋白激酶偶联受体（protein kinase linked receptor）等 3 种类型。

2. 细胞内受体能够直接传递信号或通过特定的途径传递信号　有许多细胞内受体是基因表达的调控蛋白，当这些受体与进入细胞的信号分子结合后，可以直接传递信号，即直接调控基因表达。另有一些细胞内受体可以结合细胞内产生的信号分子（如细胞应激反应中产生的细胞内信号分子），直接激活效应分子或通过一定的信号转导途径激活效应分子。

3. 膜受体的功能是识别和转换信号　膜受体有两方面的作用，一是特异性识别与结合细胞外信号分子；二是将细胞外信号转换成为能够被细胞内分子识别的信号，通过信号转导途径将信号传递至效应分子，引起细胞的应答。受体与细胞外信号分子的相互作用具有高度特异性、高亲和力、可饱和性、可逆性等特点。

（三）细胞内信号转导需要不同类型的转导分子

细胞外的信号经过受体转换进入细胞内，通过细胞内一些蛋白质和小分子活性物质进行传递，这些能够传递信号的分子称为信号转导分子（signal transducer）。细胞内参与信号转导途径的信号转导分子主要有三大类。

1. 小分子化学物质　环腺苷酸（cyclic AMP，cAMP）、环鸟苷酸（cGMP）、甘油二酯（diacylglycerol，DAG）、三磷酸肌醇（Inositol-1，4，5-triphosphate，IP_3）、磷脂酰肌醇-3,4,5-三磷酸（PIP_3）、Ca^{2+} 等小分子化学物质可以作为外源信息在细胞内的信号转导分子，称为细胞内小分子信使，或称为第二信使（second messenger）。这类信号转导分子一般具有三个特点：①不位于能量代谢途径的中心；②在细胞中的浓度或分布可以迅速改变；③作为变构效应剂作用于相应的靶分子。

2. 酶分子　许多信号转导分子都是酶。依据其转导信号的方式，作为信号转导分子的酶主要有两大类。一是催化小分子信使生成和转化的酶，如腺苷酸环化酶（adenylate cyclase，AC）、鸟苷酸环化酶（guanylate cyclase，GC）、磷脂酶 C（phospholipase C，PLC）、磷脂酶 D（PLC）等；二是蛋白激酶（protein kinase），作为信号转导分子的蛋白激酶主要是蛋白酪氨酸激酶（protein tyrosine kinase，PTK）和蛋白丝氨酸/苏氨酸激酶。

蛋白酪氨酸激酶又可分为3种类型：①蛋白酪氨酸激酶受体：此类受体胞外部分为配体结合区，中间有跨膜区，胞内部分含有蛋白酪氨酸激酶的催化结构域，此激酶活性区可催化受体胞内部分的酪氨酸磷酸化（自身磷酸化作用）；②位于细胞质部分的蛋白酪氨酸激酶：这些激酶都是信号转导途径中的中间转导分子；③核内蛋白酪氨酸激酶：这些激酶被上游信号转导分子激活后，直接作用于效应蛋白，引起细胞应答反应。

3. 调节蛋白　信号转导途径中有许多信号转导分子是没有酶活性的蛋白质，它们通过分子间的相互作用被激活、或激活下游分子。这些信号转导分子主要包括G蛋白和接头蛋白。

（1）G蛋白：GTP/GDP结合蛋白简称G蛋白（G protein）。参与细胞信号转导的G蛋白有两大类：①G蛋白三聚体：G蛋白偶联受体主要是与这一类G蛋白相互作用。此类G蛋白以α亚基、β亚基、γ亚基三聚体的形式存在于细胞质膜内侧；②小G蛋白（small G protein）：此类G蛋白为单体，分子量只有21 kDa。G蛋白结合GTP时成为活化形式，当其将GTP水解成为GDP时（G蛋白自身具有GTP酶活性）则回到非活化状态，活化的G蛋白通过变构调节作用激活下游信号转导分子。

（2）接头蛋白：接头蛋白亦称衔接蛋白（adapter proteins），可介导蛋白质信号转导分子之间或蛋白质信号转导分子与脂类分子间相互作用，通过变构效应激活下游分子。接头蛋白之间或接头蛋白与其他信号转导蛋白之间的相互作用是通过特定的结构实现的：一个蛋白分子具有特殊的结合位点（通常是由3～6个氨基酸残基组成），另一蛋白分子具有与之特异结合的结构域（大约由50～100个氨基酸残基构成）。目前已知的结构域已有近50种，以下仅举数例：

SH2结构域（Src homology 2 domain）：SH2结构域识别和结合SH2结合位点，该位点由蛋白分子中的磷酸化酪氨酸及相邻的3～6个氨基酸残基所构成。此位点的序列通常为pYxxγ，其中pY代表磷酸化酪氨酸，γ代表疏水氨基酸。不同的氨基酸组成可形成不同的SH2结合位点，与不同蛋白质分子的SH2结构域结合。

SH3结构域（Src homology 3 domain）：SH3结构域识别和结合蛋白分子中富含脯氨酸的序列（R/KxxPxxP或PxxPxR/K），其亲和力亦与脯氨酸残基及邻近氨基酸残基组成相关。

PH结构域（pleckstrin homology domain）：PH结构域可以与磷脂类分子PIP_2、PIP_3等结合，介导蛋白质定位于膜。此外，一些蛋白分子，如PKC和G蛋白的βγ亚单位也可以与PH结构域结合。

PTB结构域（phosphotyrosine binding domain）：与SH2一样，PTB结构域也可以识别一些含磷酸化酪氨酸的位点，但其氨基酸序列与SH2结合位点有所差别。

上述结构域的主要特点见图8-1。除这些结构域之外，许多接头蛋白也含有富含脯氨酸序列区，可结合其他信号蛋白分子中的SH3区，或者含有酪氨酸序列区，在酪氨酸磷酸化后能与其他信号蛋白的SH2或PTB区发生结合。

二、信号转导分子以两种基本方式向下游传递信号

每一条信号转导途径都是由多种信号转导分子组成的，不同分子间有序地依次进行相互作用，上游分子引起下游分子的数量、分布或活性状态变化，从而使信号向下游传递。信号转导分子相互作用的机制构成了信号转导的基本机制。

（一）通过改变小分子的浓度和细胞内定位而传递信号

1. 酶促反应生成小分子信使　具有酶活性的信号转导分子被其上游分子激活后，催

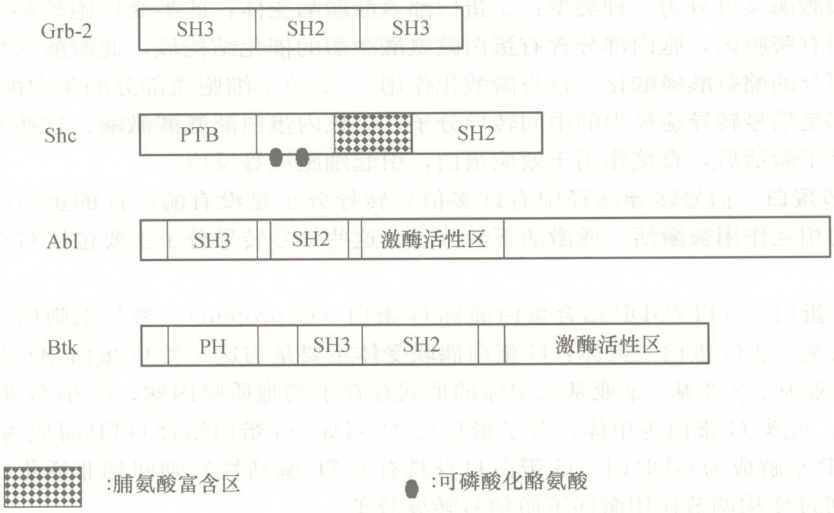


图 8-1　蛋白质相互作用结构域的特点

蛋白质相互作用的结构域的主要特点：①一个信号分子可以含有两种以上的结构域，因此可以同时与两种以上的其他信号转导分子相结合；②同一类结构域可存在于多种不同的信号转导分子中，这些结构域的一级结构不完全一致，因此对所结合的信号分子具有选择性，这是保证信号分子相互作用具有特异性的基础；③这些结构域本身均无催化活性

化特定的化学反应而产生下游分子（小分子信使），使其数量迅速增加。如腺苷酸环化酶被激活后催化 ATP 生成 cAMP，鸟苷酸环化酶激活后催化 GTP 生成 cGMP，磷脂酶 C（PLC）激活后催化 PIP_2（磷脂酰肌醇-4,5-二磷酸）水解生成 IP_3 和 DAG。

2. 调控离子通道可改变小分子的浓度或细胞内的定位分布　受体、蛋白激酶、IP_3 等信号分子可作用于钙离子通道，使其开放，增加细胞质中的钙离子浓度，使钙离子在细胞内的浓度和分布发生改变，介导信号向下游传递。

（二）上游分子通过变构激活下游分子而传递信号

许多蛋白质信号转导分子可通过构象变化而在有活性和无活性两种状态之间进行转换，这些分子经上游分子作用而激活，激活后又可激活其下游分子。信号转导途径中上游分子变构激活下游分子的模式主要有以下几种。

1. 配体结合并激活受体　脂溶性信号分子进入细胞后，作为配体与受体结合，使受体变构，暴露出 DNA 结合部位，从而能够与 DNA 结合（图 8-2）。而水溶性信号分子与细胞膜表面受体结合后，也是通过诱导受体构象改变而将其激活。

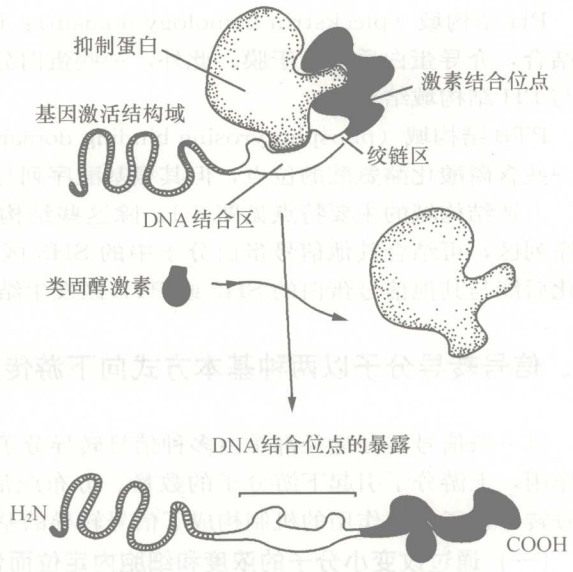


图 8-2　配体引起受体蛋白质变构

2. 酶分子通过共价修饰变构激活下游转导分子　许多蛋白质信号转导分子可经共价修饰而发生构象改变，从而在有活性和无活性两种状态之间转换。蛋白激酶将磷酸基团转移到这些蛋白分子中的丝氨酸、苏氨酸或酪

氨酸残基的羟基上（磷酸化），使其构象改变而引起活性的改变。

3. 小分子信使结合并激活下游分子　小分子信使都没有酶活性，一般是直接结合其下游分子，引起下游蛋白分子变构而将其激活。如 cAMP 激活蛋白激酶 A（PKA）（图 8-3），cGMP 激活 PKG（PKG）等。

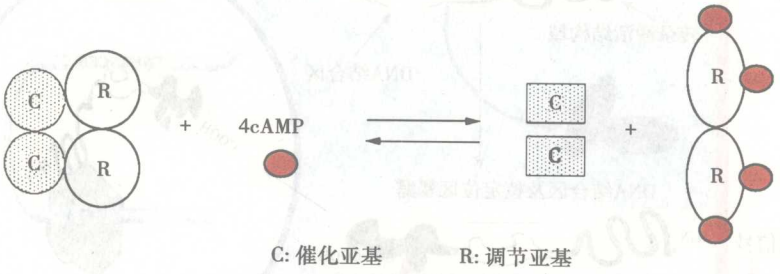


图 8-3　cAMP 诱导 PKA 变构激活

4. 上游蛋白质分子结合激活下游蛋白质分子　许多蛋白质信号转导分子也可通过变构调节的方式结合并激活下游分子。蛋白质分子间相互作用的重要结构基础之一是特定的结构域和与之对应的结合位点。在这种调节作用中，上游分子需经激活后才能与下游分子相互作用，目前已知至少有两种方式：①上游分子激活后形成或暴露出相互作用部分：如 G 蛋白都是经上游分子激活后，释放 GDP、结合 GTP，才能与下游效应分子相互作用。②共价修饰后产生特定结合位点：某些蛋白质磷酸化后产生结合位点，从而能够与下游蛋白质分子中的相应结构域结合。如上游信号转导分子经修饰后形成 SH2 结合位点，即可与含有 SH2 结构域的下游分子结合，并通过变构调节作用而将其激活。

（三）信号转导体的形成可促进信号的转导

在信号转导过程中，上游分子与下游分子的相互作用并不总是一对一的进行。有时需要几种分子结合在一起，才能发生相互作用，或产生更强的作用。几种信号转导分子结合形成的复合体称为信号转导复合物（signaling complex）或信号转导体（signalsome）。接头蛋白的功能是特异性地介导信号蛋白分子之间或信号蛋白与脂类分子之间的相互结合，因此，接头蛋白往往是信号转导体的核心成分。

第二节　不同受体介导的信号转导途径

不同信号转导分子的特定组合及有序的相互作用，构成不同的信号转导途径。因此，对信号转导的了解，关键是各种信号转导途径中信号转导分子的基本组成、相互作用及引起的细胞应答。

大部分信号转导途径都是从细胞外信号分子与细胞受体的相互作用开始。每种类型受体都有许多种，同一类型的受体所介导的信号转导途径有许多共同之处。

一、细胞内受体直接传递信号

固醇类激素及其他脂溶性化学信号的受体存在于细胞内。激素直接进入细胞内，与受体相互作用而引起细胞功能改变（图 8-4）。

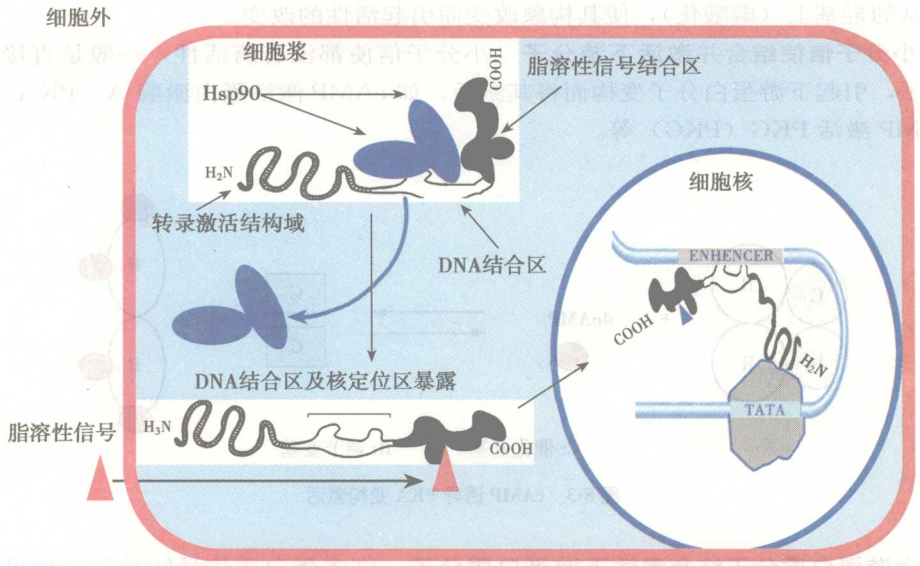


图 8-4　细胞内受体介导的信号转导

脂溶性激素受体存在于细胞内，它们通常都是转录因子，分子中具有锌指结构作为其 DNA 结合区。在没有激素作用时，受体与抑制蛋白（热休克蛋白）形成复合物，因此阻止了受体向细胞核的移动及其与 DNA 的结合。当激素与受体结合时，受体构象发生变化，导致热休克蛋白与其解聚，暴露出受体的核内转移部位及 DNA 结合部位，激素-受体复合物向核内转移，并结合在特异基因的激素反应元件上，诱导相应的基因表达，引起细胞功能改变

二、离子通道型受体通过控制离子流动而传递信号

离子通道型受体是一类自身为离子通道的受体，通道的开放或关闭直接受化学配体（主要为神经递质）的控制。离子通道是由寡聚体形成的孔道，每个单体都有 4 个跨膜区段。图 8-5 显示了作为离子通道受体典型代表的乙酰胆碱受体的结构模式。乙酰胆碱结合受体并改变其构象，从而控制离子通道的开放（图 8-6）。

离子通道型受体可以是阳离子通道，如乙酰胆碱、谷氨酸和五羟色胺的受体；也可以是阴离子通道，如甘氨酸和 γ-氨基丁酸的受体。离子通道受体信号转导的作用是改变细胞膜电位，这类受体引起的细胞应答主要是去极化与超极化。

三、G 蛋白和小分子信使是 G 蛋白偶联受体介导的信号转导途径的主要特征

神经递质、肽类激素、趋化因子、外源性刺激（味、光等）等细胞外信号可以通过 G 蛋白偶联受体（GPCR）传递信号。GPCR 在结构上均为单体蛋白，氨基端位于细胞膜外表面，羧基端在胞膜内侧，完整的肽链反复跨膜七次，因此又称为七跨膜受体。由于肽链反复跨膜，在膜外侧和膜内侧形成了几个环状结构，它们分别负责接受外源信号（化学、物理信号）的刺激和细胞内的信号传递，其细胞质部分可以与三聚体 G 蛋白相互作用。此类受体通过 G 蛋白向下游传递信号，因此称为 G 蛋白偶联受体。

（一）G 蛋白偶联受体介导的信号转导途径具有相同的基本模式

三聚体 G 蛋白是由 α、β、γ 亚基组成。目前已经发现了 20 余种 α 亚基，β 和 γ 亚基亦有数种，从而可以产生多种不同组合的三聚体，形成功能不同的 G 蛋白。α 亚基具有多个功能位点，其中包括可与受体结合并受其活化调节的部位、与 βγ 亚基相结合的部位、

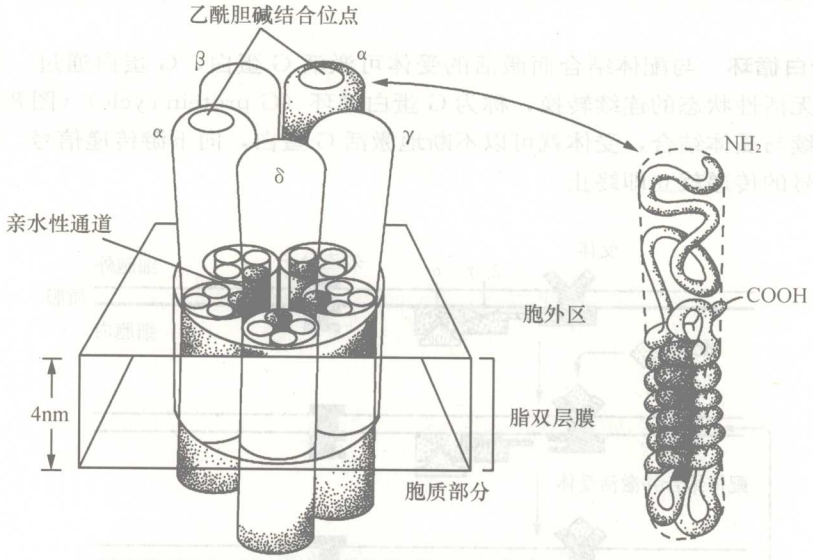


图 8-5　乙酰胆碱受体的结构模式图

乙酰胆碱受体由 5 个同源性很高的亚基构成，包括 2 个 α 亚基，1 个 β 亚基，1 个 γ 亚基和 1 个 δ 亚基。每一个亚基都是一个四次跨膜蛋白，跨膜部分为四个 α 螺旋结构，其中一个是亲水性 α 螺旋（含较多的极性氨基酸），五个亚基的亲水性 α 螺旋共同在膜中形成一个亲水性的通道。乙酰胆碱的结合部位位于 α 亚基上

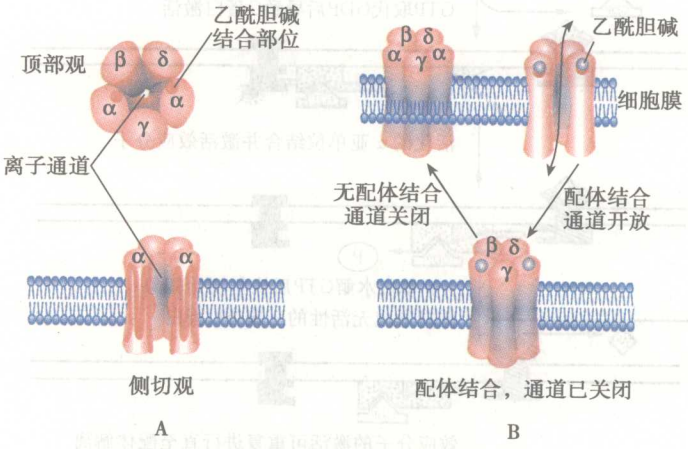


图 8-6　乙酰胆碱受体结合配体时的构象变化

乙酰胆碱受体可以三种构象存在。两分子乙酰胆碱的结合可以使之处于通道开放构象，但即使有乙酰胆碱的结合，通道开放构象持续的时间仍十分短暂，在几十毫微秒内又回到关闭状态，然后乙酰胆碱与之解离，受体恢复到初始状态，做好重新接受配体的准备

GDP 或 GTP 结合部位以及与下游效应分子相互作用的部位等，α 亚基还具有 GTP 酶活性。在细胞内，β 和 γ 亚基形成紧密结合的二聚体，其主要作用是与 α 亚基形成复合体并定位于质膜内侧，但也可直接作用于下游效应分子。

不同的 G 蛋白（不同的 αβγ 组合）可与不同的下游分子组成信号转导途径。GPCR 介导的信号传递可通过不同的途径产生不同的效应，但信号转导途径的基本模式大致相同，主要包括以下几个步骤或阶段。

1. 配体结合并激活受体　细胞外信号分子结合受体，通过变构调节作用将其激活。

当细胞外信号分子的浓度降至一定水平，即与受体解离，受体恢复到无活性状态，停止传递信号。

2. G 蛋白循环　与配体结合而激活的受体可激活 G 蛋白，G 蛋白通过一定的机制进行有活性和无活性状态的连续转换，称为 G 蛋白循环（G protein cycle）（图 8-7）。如果细胞外配体持续与受体结合，受体就可以不断地激活 G 蛋白，向下游传递信号。如果外源配体消失，信号的传递就立即终止。

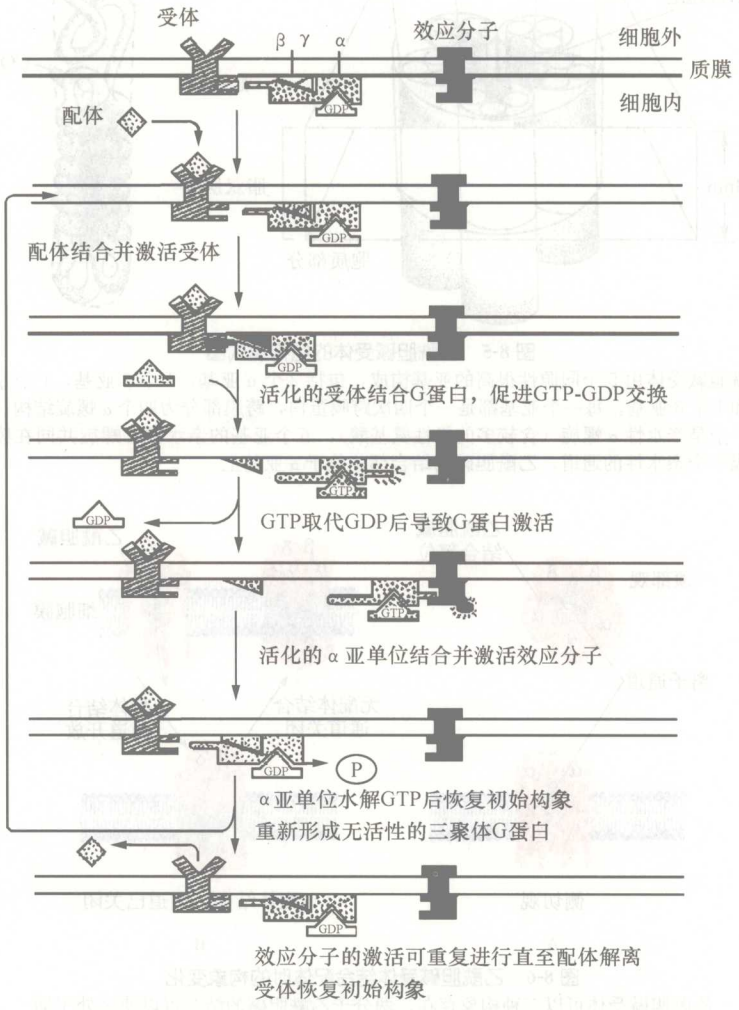


图 8-7　G 蛋白循环

配体激活的受体与 C 蛋白结合，使之发生构象改变，α 亚基与 GDP 的亲和力下降，因而释放 GDP，随即与 GTP 结合。α 亚基结合了 GTP 后即与 βγ 亚基解离，成为活化状态的 α 亚基，能够结合并激活下游的效应分子。而下游分子则可激活 α 亚基的 GTP 酶活性，将 GTP 水解成 GDP，α 亚基又恢复到原来的构象，从而与效应分子解离，重新与 βγ 亚基结合形成三聚体，回到静止状态

3. G 蛋白激活下游效应分子　活化的 G 蛋白可激活其下游的效应分子。不同的 α 亚基激活不同的效应分子，如 AC、PLC 等效应分子都是由不同的 G 蛋白所激活（表 8-1）。有的 α 亚基可以激活腺苷酸环化酶，称为 αs（s 代表 stimulate）；有的 α 亚基可以抑制腺苷酸环化酶，称为 αi（i 代表 inhibit）。

4. 小分子信使的产生或分布变化　G 蛋白的效应分子向下游传递信号的主要方式是催

表 8-1　哺乳动物细胞中的 Gα 亚基种类及效应

G 蛋白种类	效应	产生的第二信使	第二信使的靶分子
αs	AC 活化↑	cAMP↑	PKA 活性↑
αi	AC 活化↓	cAMP↓	PKA 活性↓
αq	PLC 活化↑	Ca^{2+}、IP_3、DAG↑	PKC 活化↑
α_t	cGMP-PDE 活性↑	cGMP↓	Na^+通道关闭

PDE：phosphodiesterase，磷酸二酯酶

化产生小分子信使，如 AC 催化产生 cAMP，PLC 催化产生 DG 和 IP_3。有些效应分子可以通过对离子通道的调节改变 Ca^{2+} 在细胞内的分布，其效应与 IP_3 的效应相似。

5. 小分子信使激活蛋白激酶　小分子信使作用于相应的靶分子（主要是蛋白激酶），使之构象改变而激活。

6. 蛋白激酶激活效应蛋白　蛋白激酶通过磷酸化作用激活一些与代谢相关的酶、与基因表达相关的转录因子，以及一些与细胞运动相关的蛋白，从而产生各种细胞应答反应。

（二）不同 G 蛋白偶联受体可通过不同途径传递信号

不同的细胞外信号分子与其受体结合后，通过 G 蛋白传递信号，但传入细胞内的信号并不一样。这是因为不同的 G 蛋白与不同的下游分子组成了不同的信号转导途径。有几条途径是较常见的，本节主要介绍其中 3 条途径（图 8-8）。

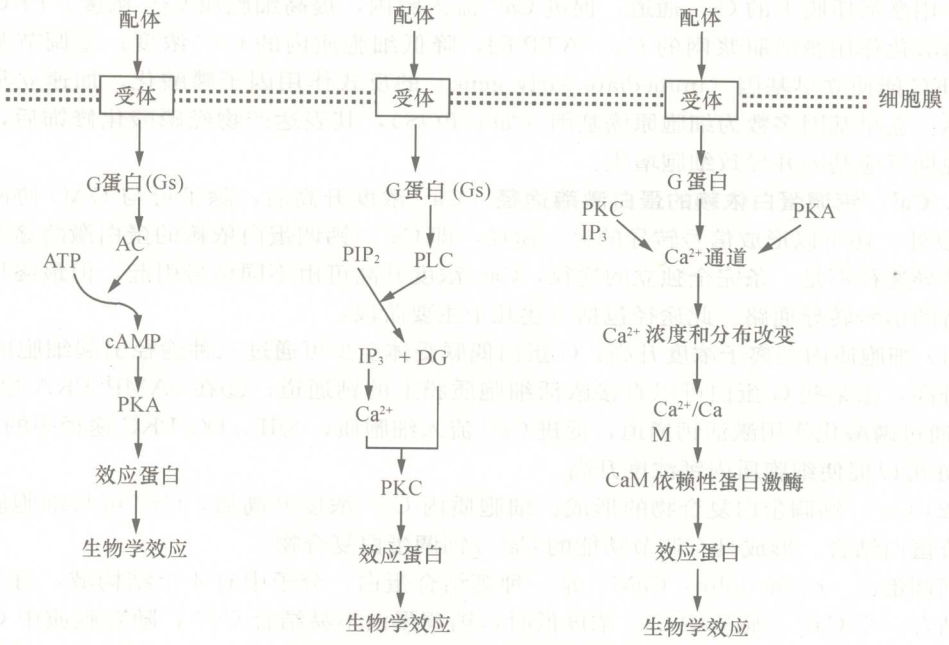


图 8-8　G 蛋白偶联受体介导的信号转导

1. cAMP-PKA 途径　此类信号转导途径中的细胞外信号分子包括胰高血糖素、肾上腺素、促肾上腺皮质激素等。激素与受体结合后激活 G 蛋白，释放活化的 α 亚基，α 亚基激活腺苷酸环化酶（AC）。AC 催化产生小分子信使 cAMP。cAMP 结合 PKA，通过变构调节作用激活 PKA（图 8-3）。

PKA 通过磷酸化作用激活或抑制各种效应蛋白，继续传递信号。

（1）调节代谢：PKA 可通过磷酸化作用调节十几种关键酶的活性，对不同的代谢途径发挥调节作用，如激活磷酸化酶 b 激酶、激素敏感脂肪酶、胆固醇酯酶，促进糖原、脂肪、胆固醇的分解代谢；抑制乙酰 CoA 羧化酶、糖原合成酶 I 等，抑制脂肪合成和糖原合成。

（2）调节基因表达：PKA 可通过磷酸化修饰作用激活反式作用因子，进而激活基因表达。如激活的 PKA 进入细胞核，使 cAMP 反应元件结合蛋白（cAMP response element binding protein，CREB）磷酸化。磷酸化的 CREB 结合于 cAMP 反应元件（CRE），并与 CREB 结合蛋白（CREB-binding protein，CBP）结合。与 CREB 结合后的 CBP 作用于通用转录因子（包括 TFIIB），促进 TFIIB 等通用转录因子与启动子结合，激活基因的表达。

（3）调节细胞极性：PKA 亦可通过磷酸化作用激活离子通道，调节细胞膜电位。

2. IP_3/DAG-PKC 途径 此类信号转导途径中的细胞外信号分子包括促甲状腺素释放激素、去甲肾上腺素、抗利尿素等。与激素结合的受体激活 G 蛋白，释放活化的 α 亚基，α 亚基激活 PLC。PLC 水解膜组分 PIP_2，生成 DAG 和 IP_3。IP_3 从膜上扩散到细胞质中，与内质网和肌浆网上的受体结合，促进这些钙储库内的 Ca^{2+} 迅速释放，使细胞质内的 Ca^{2+} 浓度升高。Ca^{2+} 与细胞质内的 PKC 结合并聚集至质膜。DAG 生成后仍留在质膜上，当结合有 Ca^{2+} 的 PKC 聚集到质膜时，DAG、磷脂酰丝氨酸与 Ca^{2+} 共同作用于 PKC 的调节结构域，使 PKC 变构而暴露出活性中心。

PKC 通过磷酸化作用激活各种功能蛋白，调节细胞生理活动、激活基因表达。其作用有：①调节细胞生理活动：PKC 的底物蛋白包括质膜受体、膜蛋白和多种酶。通过磷酸化作用激活质膜上的 Ca^{2+} 通道，促进 Ca^{2+} 流入胞内，提高细胞质 Ca^{2+} 浓度。PKC 也能通过磷酸化作用激活肌浆网的 Ca^{2+}-ATP 酶，降低细胞质内的 Ca^{2+} 浓度；②调节基因表达：PKC 能使立早基因（immediate-early gene）的反式作用因子磷酸化，加速立早基因的表达。立早基因多数为细胞原癌基因（如 *c-FOS*），其表达产物经磷酸化修饰后，最终活化晚期反应基因并导致细胞增生。

3. Ca^{2+}/钙调蛋白依赖的蛋白激酶途径 Ca^{2+} 浓度升高后，除了可与 DAG 协同激活 PKC 以外，还可以形成信号转导的另一途径，即 Ca^{2+}/钙调蛋白依赖的蛋白激酶途径。此信号转导途径不是一条完全独立的途径，Ca^{2+} 浓度升高可由不同信号引起，但最终形成一条独特的信号转导通路。此途径包括下述几个主要阶段：

（1）细胞质内钙离子浓度升高：G 蛋白偶联受体至少可通过三种途径引起细胞内 Ca^{2+} 浓度升高：①某些 G 蛋白可以直接激活细胞质膜上的钙通道；②在 cAMP/PKA 途径中，PKA 通过磷酸化作用激活钙通道，促进 Ca^{2+} 流入细胞质；③IP_3/DG-PKC 途径中的 IP_3 和 PKC 亦可以促使细胞质内钙浓度升高。

（2）Ca^{2+}/钙调蛋白复合物的形成：细胞质内 Ca^{2+} 浓度升高后，Ca^{2+} 可与细胞质内的钙调节蛋白结合，形成具有调节功能的 Ca^{2+}/钙调蛋白复合物。

钙调蛋白（calmodulin，CaM）是一种钙结合蛋白，分子中有 4 个结构域，每个结构域可结合一个 Ca^{2+}。胞质中 Ca^{2+} 浓度低时，钙调蛋白不易结合 Ca^{2+}；随着胞质中 Ca^{2+} 浓度增高，钙调蛋白可结合不同数量的 Ca^{2+}，形成不同构象的 Ca^{2+}/CaM 复合物。钙调蛋白本身无活性，形成 Ca^{2+}/CaM 复合物后则具有调节功能。

（3）Ca^{2+}/CaM 复合物激活蛋白激酶：Ca^{2+}/CaM 复合物的下游信号转导分子是一些蛋白激酶。这些蛋白激酶能催化许多效应蛋白的磷酸化，它们的共同特点是可被 Ca^{2+}/CaM 复合物激活，因而统称为钙调蛋白依赖性蛋白激酶。

（4）钙调蛋白依赖性激酶激活效应蛋白：钙调蛋白依赖性激酶能使底物蛋白在 Ser 和 Thr 残基上发生磷酸化，属于蛋白丝/苏氨酸激酶，如肌球蛋白轻链激酶（MLCK）、

磷酸化酶激酶（PhK）、钙调蛋白依赖性激酶（Cal-PK）Ⅰ、Ⅱ、Ⅲ等。这些激酶可激活各种效应蛋白，可在收缩和运动、物质代谢、神经递质的合成、细胞分泌和分裂等多种生理过程中起作用。如 Cal-PK Ⅱ可修饰激活突触蛋白 I、酪氨酸羟化酶、色氨酸羟化酶、骨骼肌糖原合成酶等，参与神经递质的合成与释放以及糖代谢等多种细胞功能的调节。

四、蛋白激酶偶联受体介导的信号转导途径主要涉及蛋白质的相互作用

蛋白激酶偶联受体主要是生长因子和细胞因子的受体，组成受体的每一条肽链只跨膜1次，因而又称为单跨膜受体。此类受体介导的信号转导主要是调节蛋白质的功能和表达水平、调节细胞增殖和分化。蛋白激酶偶联受体主要通过蛋白质分子的相互作用激活细胞内蛋白激酶。

（一）蛋白激酶偶联受体介导的信号转导途径也具有相同的基本模式

蛋白激酶偶联受体介导的信号转导途径较复杂。细胞内的蛋白激酶有许多种，不同蛋白激酶的组合形成不同的信号转导途径，具体作用模式也有差别。但蛋白激酶偶联受体介导的信号转导途径的基本模式大致相同，主要包括以下几个阶段：

（1）胞外信号分子与受体结合。

（2）第一个蛋白激酶被激活。这一步反应是“蛋白激酶偶联受体”名称的由来。但是，“偶联”却有两种形式。有的受体本身具有蛋白激酶活性，此步骤是激活受体胞内结构域的蛋白激酶活性。有些受体本身没有蛋白激酶活性，此步骤是受体通过蛋白质-蛋白质相互作用激活某种蛋白激酶。

（3）通过蛋白质-蛋白质相互作用或蛋白激酶的磷酸化修饰作用激活下游信号转导分子、传递信号。但最终仍是激活一些特定的蛋白激酶。

（4）蛋白激酶通过磷酸化修饰激活代谢途径中的关键酶、反式作用因子等，影响代谢途径、基因表达、细胞运动、细胞增殖等。

（二）几种常见的蛋白激酶偶联受体介导的信号转导途径

目前已发现的蛋白激酶偶联受体介导的信号转导途径有十几条，本节介绍几条较常见的途径。

1. MAPK 途径　以丝裂原激活的蛋白激酶（mitogen-activated protein kinase，MAPK）为代表的信号转导途径称为 MAPK 途径（图 8-9）。在不同的细胞中，该途径的成员组成及诱导的细胞应答有所不同。其中了解最清楚的是 Ras/MAPK 途径。

Ras/MAPK 途径转导生长因子（如表皮生长因子 EGF）信号，其基本过程是：①受体与配体结合后形成二聚体，激活受体的蛋白激酶活性；②受体自身酪氨酸残基磷酸化，形成 SH2 结合位点，从而能够结合接头蛋白 Grb2（图 8-1）；③Grb2 的两个 SH3 结构与 SOS 分子中的富含脯氨酸序列结合，将 SOS 活化；④活化的 SOS 结合 Ras 蛋白，促进 Ras 释放 GDP、结合 GTP；⑤活化的 Ras 蛋白（Ras-GTP）可激活 MAPK 激酶的激酶（MAPKKK），活化的 MAPKKK 可磷酸化 MAPK 激酶（MAPKK）而将其激活，活化的 MAPKK 将 MAPK 磷酸化而激活；⑥活化的 MAPK 可以转位至细胞核内，通过磷酸化作用激活多种效应蛋白，从而使细胞对外来信号产生生物学应答。

MAPKKK 有许多种，如 Raf、MEKK 家族（MEKK1～4）、MLK 家族（MLK1～3）、TAK、ASK 家族（ASK1～3）等。MAPKK 亦有多种，如 MEK 家族、MKK 家族等。MAPK 至少有 12 种，分属于 ERK 家族、p38 MAPK 家族、JNK 家族。这 3 条信号转导通路的组成和信号转导的细胞内效应见图 8-9。

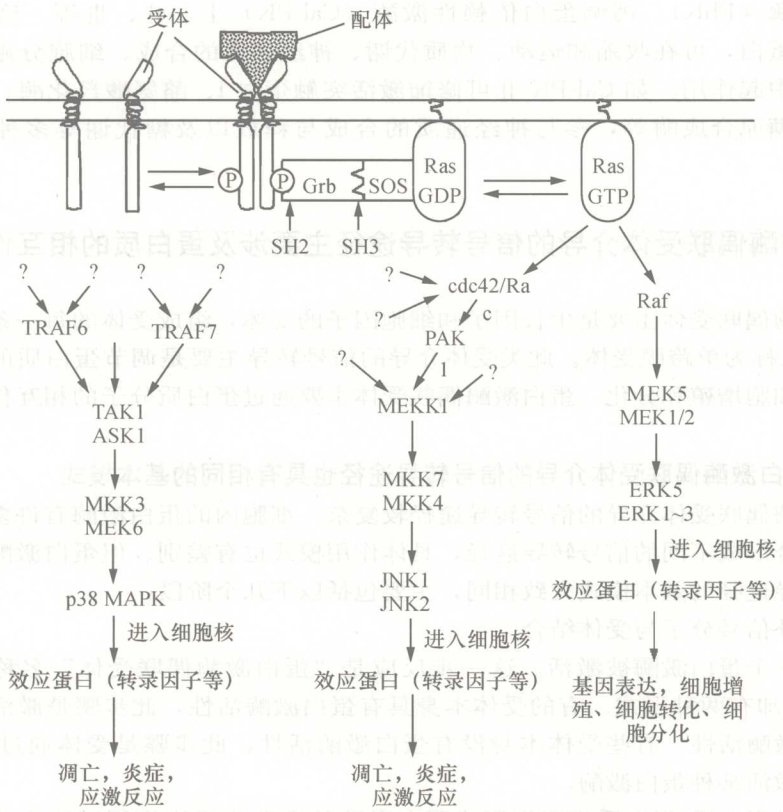


图 8-9　蛋白激酶偶联受体介导信号转导的 MAPK 途径

2. JAK-STAT 途径　许多细胞因子受体自身没有激酶结构域，与细胞因子结合后，受体通过蛋白酪氨酸激酶 JAK（Janus kinase，JAK）的作用使受体自身和胞内底物磷酸化。JAK 的底物是信号转导子/转录活化子（signal transducer and activator of transcription，STAT），二者所构成的 JAK-STAT 通路是细胞因子信息内传最重要的一条通路。

不同的受体利用不同的 JAK 和 STAT 分子。例如，IFN-γ 激活 JAK-STAT1 途径的主要步骤是：①IFN-γ 结合受体并诱导其形成同型二聚体；②受体与 JAK 结合并将其激活，JAK1 和 JAK2 成为相邻蛋白，从而相互磷酸化，并将受体磷酸化；③JAK 将 STAT1 磷酸化，磷酸化的 STAT 分子彼此间通过 SH2 结合位点和 SH2 结构域结合而二聚化，并从受体复合物中解离；④磷酸化的 STAT 同源二聚体转移到核内，调控基因的转录。

3. Smad 途径　转化生长因子 β（transform growth factor β，TGF-β）基因超家族的受体介导的信号转导途径中最重要的信号转导分子是 Smad 分子，因而此途径称为 Smad 途径。TGF-β 超家族受体主要有Ⅰ型和Ⅱ型，激活后都具有丝氨酸/苏氨酸蛋白激酶活性，可将 Smad 磷酸化。

TGF-β 信号转导的基本步骤是：①TGF-β 同时结合 2 个Ⅰ型受体和 2 个Ⅱ型受体，形成异源四联复合物，Ⅱ型受体被激活，其激酶活性将Ⅰ型受体磷酸化并活化；②膜相关蛋白 SARA（Smad anchor for receptor activation）结合 Smad2 和 Smad3，并提呈给活化的Ⅰ型受体；③受体将 Smad2 和 Smad3 磷酸化；④磷酸化的 Smad2 和 Smad3 与 Smad4 形成三聚体转移至细胞核内，结合于 Smad 结合元件，调节基因表达。

4. PI3K/PKB 途径　磷脂酰肌醇-3-激酶（phosphoinositide 3-kinase，PI3K）和蛋白

激酶 B（protein kinase B，PKB）共同构成一条重要的信号转导通路。PKB 是原癌基因 *c-akt* 的产物，故又称为 Akt。配体与受体结合后，PI3K 的 p85 亚单位与活化的受体结合，p110 亚单位被受体磷酸化而活化。PI3K 可催化 PIP_3 的产生。PIP_3 产生后，通过结合 PKB 的 PH 结构域，将其锚定于质膜而活化。PKB 可磷酸化多种蛋白，介导代谢调节、细胞存活等效应。

第三节　信号转导的基本规律和复杂性

一、各种信号转导机制具有共同的基本规律

（一）信号的传递和终止涉及许多双向反应

信号的传递和终止实际上就是信号转导分子的数量、分布、活性转换的双向反应。如 AC 催化生成 cAMP 而传递信号，而磷酸二酯酶则将 cAMP 迅速水解为 5′-AMP 而终止信号传递。以钙离子为细胞内信使时，Ca^{2+} 可以从其贮存部位迅速释放，然后又通过细胞 Ca^{2+} 泵作用迅速恢复初始状态。PLC 催化 PIP_2 分解成 DAG 和 IP_3 而传递信号，DAG 激酶和磷酸酶分别催化 DAG 和 IP_3 转化而重新合成 PIP_2。对于蛋白质信号转导分子，则是通过与上、下游分子的迅速结合与解离而传递信号或终止信号传递，或者通过磷酸化作用和去磷酸化作用在活性状态和无活性状态之间转换而传递信号或终止信号传递。

（二）细胞信号在转导过程中被逐级放大

细胞在对外源信号进行转换和传递时，大都具有逐级将信号加以放大的作用。G 蛋白偶联受体介导的信号转导过程和蛋白激酶偶联受体介导的 MAPK 途径都是典型的级联反应过程。

（三）细胞信号转导途径既有通用性又有特异性

细胞内许多信号转导分子和信号转导途径常常为不同的受体共用，而不是每一个受体都有专用的分子和途径。换句话讲，细胞的信号转导系统对不同的受体具有通用性。例如，许多受体可以通过 Ras→MAPK 途径传递信号。信号转导途径的通用性使得细胞内有限的信号转导分子可以满足多种受体信号转导的需求。另一方面，不同的细胞具有不同的受体，而同样的受体在不同的细胞又可利用不同的信号转导途径，同一信号转导途径在不同细胞中的最终效应蛋白又有所不同。因此，细胞-受体-信号转导途径-效应蛋白可以有多种不同组合，而一种特定组合决定了一种细胞对特定的细胞外信号分子产生特异性应答。

二、一条信号转导途径中的功能分子可影响和调节其他途径

细胞内的信号转导途径并不是各自独立存在，不同途径之间存在着多种交互的联系。当一条途径中的信号转导分子对另一条途径中的信号转导分子发挥修饰作用时，可对该途径发挥调控作用。下面两个例子可使我们对此有一个初步了解。

（一）Ras/MAPK 途径可调节 Smad 途径

Ras/ERK 途径转导的信号可促进细胞增殖，而 Smad 途径转导的信号则抑制细胞增殖。对于正常上皮细胞，作为维持细胞稳态的 TGF-β 占主导地位，并对抗由生长因子经 Ras 途径激活的增殖反应。然而，当大量的生长因子（如 EGF、HGF）刺激细胞或 *RAS* 癌基因激活后，使 Ras/ERK 途径激活，活化的 ERK1/2 蛋白激酶将 Smad1/2/3 等分子磷酸化，使 Smad1/2/3 向核内聚集的能力减弱，从而抑制 Smad 传递信号的作用。此时增殖

信号成为调控细胞增殖反应的主要信号。

（二）蛋白激酶 C 可调节蛋白酪氨酸激酶系统

PKC 是肌醇磷脂系统的中心酶，它使大量的底物磷酸化而参与广泛的生理调节，如 DNA 与蛋白质的合成、细胞的生长分化、细胞分泌作用等。PKC 对蛋白酪氨酸激酶系统也有调节作用。通过磷酸化修饰 EGF 受体、白介素 2 受体、Ras、Raf-1 等，可对 Ras/MAPK 途径产生调节作用。

三、细胞信号转导复杂且具有多样性

配体-受体-信号转导分子-效应蛋白并不是以一成不变的固定组合构成信号转导途径，细胞信号转导是复杂的，且具有多样性。下面几个例子可说明这种复杂性和多样性。

（一）一种受体可激活多条信号转导途径

一种受体并非只能激活一条信号转导途径，有些受体自身磷酸化后产生多个与其他蛋白相互作用的位点，可以激活几条信号转导途径。如血小板衍生生长因子（platelet-derived growth factor，PDGF）的受体（图 8-10）激活后，可激活 Src 激酶活性、结合 Grb2 并激活 Ras、激活 PI3K、激活 PLCγ，因而同时激活多条信号转导途径而引起复杂的细胞应答反应。另一方面，一条信号转导途径也不是只能由一种受体激活。例如，有多种受体可以激活 PI3K 途径。

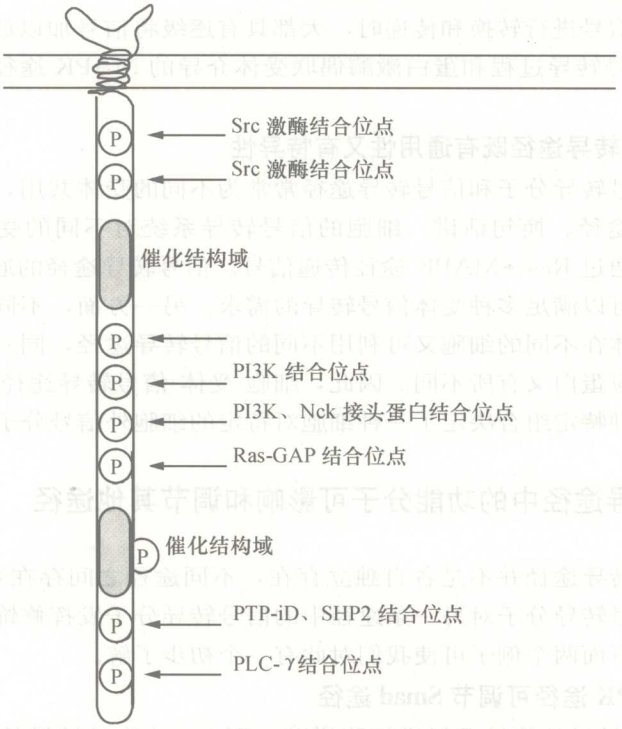


图 8-10　PDGF 受体胞内区与其他蛋白质相互作用位点和激酶催化结构域

（二）一种信号转导分子不一定只参与一条途径的信号转导

G 蛋白偶联受体（GPCRs）主要是通过激活 AC 而参与代谢调节的信号转导途径，因而 GPSRs 一般是在分化成熟（完全）的组织细胞参与信号转导。但 GPCRs 在某些增殖细

胞中也可表达，在这些细胞中，G 蛋白的 βγ 二聚体可激活 Src 或 Src 样激酶（如 Fyn、Lyn 和 Yes 蛋白酪氨酸激酶），后者使 Shc 的酪氨酸残基磷酸化，形成 SH2 结合位点。活化的 Shc 与 Grb2 相互作用形成 Shc-Grb2 复合物，通过 SOS、Ras 蛋白激活 MAPK 途径，调控细胞增殖所需基因的转录。

（三）一种细胞外信号分子可通过不同信号转导途径影响不同的细胞

白介素 1β（IL-1β）是在局部和全身炎症反应中起核心作用的细胞因子。然而，由于其受体分布广泛，IL-1β 的作用并不仅限于炎症。IL-1β 可以通过 G 蛋白偶联受体（涉及 cAMP 途径、cGMP 途径、PKC 途径等）和蛋白激酶偶联受体介导的 MAPK 途径介导信号转导。近年又发现 IL-1β 可通过其他几条重要途径介导信号转导，包括：IL-1 受体相关激酶（interleukin-1 receptor-associated kinase，IRAK）途径、PI3K 途径、JAK-STAT 通路、离子通道。IL-1β 受体在多种细胞表面存在，受体可通过不同信号转导途径传递信号，它不仅可以作用于各种炎症相关细胞，还可以通过 JAK-STAT 途径作用于胰岛 B 细胞，通过激活离子通道影响神经细胞、血管平滑肌细胞、成纤维细胞、骨髓基质细胞等多种细胞的功能。

（四）不同信号转导通路可介导相同的生物学效应

趋化因子是体内一类能够诱导特定细胞趋化运动的分子。趋化因子受体是一类表达于不同类型细胞上的 G 蛋白偶联受体。然而目前的研究结果显示，趋化因子可以通过不同的信号转导途径传递信号，如激活 PKA 途径、调节细胞内 Ca^{2+} 浓度、G 蛋白 βγ 亚单位和磷酸酪氨酰肽协同作用可激活 PI3K 途径、MAPK 途径，还可以激活 JAK-STAT 途径。这些信号途径不同，但所有趋化因子诱导产生的生物学效应都是相同的：细胞趋化运动。

（五）细胞内的特殊事件也可以启动信号转导或调节信号转导

一些特殊的细胞内事件也可以在细胞内启动信号转导途径。如 DNA 损伤、ROS、低氧状态等，可通过激活特定的分子而启动信号转导。这些途径可以与细胞外信号分子共用部分转导途径、共用一些信号分子，也可以是一些特殊的途径（如凋亡信号转导途径）。有关内容将在第十七章（应激反应的分子机制）介绍。

（冯作化）

第九章　细胞增殖分化的分子机制

关键知识点

- 细胞增殖是个体生长发育的重要过程之一，也是一种重要的生命现象。细胞增殖受到生长因子的调控，涉及很多基因的表达变化和信号传导途径的介入。真核细胞的增殖经历一个称为细胞周期的过程。细胞周期蛋白和细胞周期蛋白依赖性激酶及其抑制因子等参与细胞周期的调节。
- 生长因子是一类具有调节细胞生长与分化作用的信号分子，通过其特异受体，经信号转导最终作用于与细胞增殖相关的基因，以影响细胞的生长或分化。
- 细胞分化是指胚胎细胞获得不同形态、结构和功能特征的过程。细胞分化过程的核心是基因的选择性表达。一系列的信号分子通过不同的途径决定了细胞特异性的基因表达方式。胚胎干细胞具有能够分化形成生物个体内全部不同类型细胞的能力，即“全能性”。横向分化或转分化是指一种已经分化的细胞类型不可逆地转化为另外一种正常分化的细胞类型。
- 组织再生是生物界普遍存在的现象。不同种属动物的再生潜能差异很大，同一种属动物的不同组织的再生能力也同样具有很大差异。再生医学是一门通过研究机体正常组织特征与功能、损伤修复与再生机制，进而寻找有效的生物治疗方法，促进机体自我修复与再生，或构建新的组织与器官，以改善或恢复损伤组织和器官之功能的科学。

关键词： 细胞增殖；细胞周期；生长因子；细胞分化；干细胞

细胞的增殖（proliferation）与分化（differentiation）是生物胚胎发育、生物个体生长以及生命活动的维持过程中的两个重要事件。从一个简单的受精卵到一个独特的，具有不同于其他个体的表型和基因型、能够完成复杂的生命活动的生物体，是一个非常奇妙的过程。细胞的增殖与分化在这个过程中是如何进行，又是如何被调控，一直是生命科学家为之探索的重要问题。此外，细胞的增殖与分化还是生物个体正常生理活动的基础，并且和生物体的衰老、疾病过程密切相关。

第一节　细胞增殖的分子基础

细胞增殖（cell proliferation）是由于细胞生长（cell growth）和细胞分裂（cell division）而导致的细胞数目的增加，是个体生长发育的重要过程之一，也是一种重要的生命现象。这种现象不仅存在于胚胎时期，而且在出生后仍然存在；它不仅在正常的生理条件下发生，而且也和某些疾病密切相关。细胞增殖主要受到生长因子的调控，涉及很多基因的表达变化和细胞信号传导途径的介入。任何一个环节失调都有可能导致疾病的发生。

一、细胞增殖是生长的主要因素之一

生物体的生长包括三个方面的含义：细胞数目增多（cell proliferation）、细胞体积增大（cell enlargement）和细胞外基质的合成（cell accretion）。

虽然在生长过程中，细胞数目的增多是主要的因素之一，但是在这一过程中也同时伴随着大量细胞的死亡。生长速度是由细胞死亡和细胞增殖的速率所决定。细胞体积的增大

是生长的另一种机制。例如：神经元可以通过延伸以及轴突的生长而增大；肌细胞可通过细胞质量的增加、卫星细胞与肌纤维的融合而变大。此外，细胞还可通过分泌大量的细胞基质来扩大细胞外空间，从而达到生长的目的。例如，骨组织的大部分成分就是由细胞外的骨基质组成。

二、细胞周期是细胞分裂过程中的有序事件

真核细胞的分裂经历一系列具有固定发生顺序的事件，这一过程称为细胞周期（cell cycle）。真核细胞的细胞周期分为两大时相：间期（interphase）和分裂期（metaphase，M 期）。间期又分为 G1 期、S 期和 G2 期，细胞大部分时间处于此期中。细胞在这一时相中根据细胞生长的需要，合成新的核糖体、膜、线粒体、内质网和大多数蛋白质。染色体 DNA 的复制发生在间期的特殊时期，即 S 期。S 期处于间期的中段，S 期之前为 G1 期，之后是 G2 期。G2 期之后细胞进入 M 期。

不同生物体的细胞，甚至同一生物体的不同细胞，其细胞周期可以有很大差别。有些细胞可能会缺失细胞周期中的某些阶段。多数哺乳动物细胞的细胞周期在 10～30 小时之间。迅速分裂的人源细胞经过整个细胞周期需要 24 小时，而迅速生长的酵母细胞只需 90 分钟就可以完成整个细胞周期。多细胞生物体中一些细胞可暂时离开细胞周期，从 G2 期进入静止期，即 G0 期。G0 期的细胞处于静止状态，不分裂，也不生长。G0 期持续的时间可以是几天、几周，甚至几年，但这些细胞可以在某些生长因子的作用下被重新诱导进入 G1 期。

三、细胞周期蛋白调节细胞周期中不同时相的正确转换

许多蛋白质因子参与了细胞周期的调节，以保证细胞周期中不同时相的正确转换。细胞周期蛋白（cyclin）就是其中的一类蛋白质因子。

（一）细胞周期蛋白在模式生物早期胚胎发育的研究中被发现

细胞周期蛋白是一类随着细胞周期不同时相的转换，其浓度也随之发生变化的蛋白质，这些蛋白质通过与相应的蛋白激酶的结合对细胞周期进行调控。在 20 世纪 80 年代初期，T. Hunt 等在以海胆（sea urchin）为模式生物进行早期胚胎发育的研究中发现，细胞内有一种蛋白的含量随着细胞周期的变化而发生改变。在有丝分裂的早期，这种蛋白含量迅速增加，随后急剧下降，在下一个有丝分裂的早期又形成一高峰。这种蛋白就是细胞周期蛋白，后来被命名为细胞周期蛋白 B（cyclin B）。现已发现细胞周期蛋白是一个大家族，包含有十几个成员。不同种属的细胞周期蛋白具有高度的保守性。

（二）细胞周期蛋白依赖性激酶调节细胞周期蛋白的活性

高等真核细胞细胞周期的调节依赖于一结构上相关的异二聚体蛋白激酶家族来完成。这些蛋白激酶均由调节亚单位和催化亚单位两部分组成。其调节亚单位，即为细胞周期蛋白，其浓度随细胞周期不同时相的转换而有升高和降低；其催化亚单位被称为细胞周期蛋白依赖性激酶（cyclin-dependent kinase，CDK）。

20 世纪 80 年代初，L. Hartwell 等利用遗传学的方法在研究啤酒酵母时，发现了一群基因与细胞周期的调控相关，称为细胞分裂周期基因（cell division cycle genes，CDC genes），其中的 *Cdc*28 又被称为“start”。80 年代中期，P. Nurse 等发现了 *Cdc*2 基因，*Cdc*2 基因不但控制 G2 期向 M 的转换，而且与 start 基因一样也控制着 G1 期向 S 期的转换。后来 *Cdc*2 基因被命名为 *Cdk*1。现在已发现十余种人的 CDK 分子。T. Hunt，

L. Hartwell 和 P. Nurse 等三人由于在细胞周期研究中的贡献，于 2001 年获得了诺贝尔生理学或医学奖。

CDKs 单独存在时并不表现其激酶活性。当细胞周期蛋白与其结合时才被激活（表 9-1）。这些蛋白复合物控制了细胞周期中不同时相的转换，以及在各个时相中的一些重要事件的发生，例如：在 S 期中 DNA 的合成和 M 期中有丝分裂的进行。

表 9-1　主要的细胞周期蛋白和细胞周期蛋白依赖性激酶

细胞周期蛋白	激　酶	功　能
Cyclin D	CDK2、CDK4、CDK6	G1 期进入 S 期
Cyclin E	CDK2、CDK3	G1 期进入 S 期
Cyclin A	CDK2	早 S 期 DNA 合成的起始
Cyclin A1	CDK2	早 S 期 DNA 合成的起始
Cyclin B	CDK1	G2 期进入 M 期

值得注意的是，有一些 CDKs 不仅仅与一种细胞周期蛋白结合；一些细胞周期蛋白也可以与一种以上的 CDK 结合。此外，CDK-细胞周期蛋白复合体的生物学功能并不仅仅局限于细胞周期的调节。CDK5 在有丝分裂期后的神经元中高度表达。而 CDK8 和 CDK9 似乎主要在转录调节方面起作用。还有 CDK2 和 CDK5 也在细胞凋亡中起作用。

（三）细胞周期蛋白的泛素化降解保证周期进程

细胞周期蛋白的含量在有丝分裂晚期急剧下降，说明细胞内存在着降解这些蛋白分子的机制。目前已知，细胞周期蛋白合成后，在其 N 末端可以发生多泛素化（polyubiquitination），即共价结合多个泛素（ubiquitin）。泛素是由 76 个氨基酸组成的蛋白分子，一些蛋白的多泛素化可导致其在细胞的蛋白酶体（proteosomes）内迅速降解。

四、不同时相转换的关卡是细胞周期调控的关键环节

细胞周期中不同时相转换的调控十分精确。一个时相结束，下一个方能开始。在整个细胞周期中有三个关键的转折点或关卡（checkpoint）：G1-S 期转折点，在酵母中称为起点（start），在哺乳动物中称为限制点（restriction point）；G2-M 期转折点，决定有丝分裂的开始；M 后期-胞质分裂完成细胞分裂。这些过程都是不可逆的，因此也就决定了细胞周期进程的单向性。细胞周期中，在不同的时相细胞周期蛋白-CDKs 复合体有所不同。它们顺序产生、降解，协调一致地控制着细胞周期的进行。

（一）G1 到 S 期的转折意味着细胞进入 DNA 合成期

细胞周期中由 G1 向 S 期的转折至关重要。细胞周期一旦通过起点或限制点，就意味着从静止状态进入了 DNA 合成期。

1. Cyclin D 的作用　静止状态的细胞不表达 cyclin D。当细胞受到生长因子等刺激时，首先表达 cyclin D，cyclin D 结合 CDK4 和 CDK6，并通过磷酸化使它们活化。Cyclin D 有三种亚型（D1、D2、D3）存在，它们的半衰期都很短（30 min 左右），只要撤除生长因子，cyclin D 的水平立即下降。

2. Cyclin E 的作用　Cyclin E 的表达晚于 cyclin D，其表达在 G1 晚期达到高峰。Cyclin E 可以与 CDK2 和 CDK3 结合，通过磷酸化使它们活化。CDK2 是调节 DNA 复制开始的关键酶。CDK2 也能够与在 S 期合成的 cyclin A 及 A1 结合并使其激活，共同在 G1 到 S 期的转折中发挥作用。

3. 转录因子 E2F 的活化　Cyclin D 或 cyclin E 与 CDK 形成的复合物可以活化转录因

子 E2F，促进 DNA 合成相关基因的表达，进而促进 DNA 的合成。在 G1 早期细胞中，E2F 是与抑癌基因 Rb 蛋白（见第十一章）相结合而处于失活状态的。Cyclin-CDK 复合物使 Rb 蛋白发生磷酸化后，E2F 即可游离出来并被活化。游离的 E2F 又可以促进其自身、CDK2、cyclin E 等的合成，进一步促进 Rb 蛋白磷酸化，释放更多的 E2F，形成一种正反馈调节，最终使细胞通过 G1-S 期转折点。Rb 蛋白在随后的 S 期、G2 期和 M 期将始终保持磷酸化状态，只有当细胞结束有丝分裂再次进入 G1 期或 G0 期时，cyclin-CDK 复合体活性下降，低磷酸化的 Rb 蛋白才又可以与 E2F 结合并抑制其活性。

4. Cyclin A 的作用　为保证细胞分裂所产生的两个子细胞含有同样的基因组，每一个 DNA 复制起点在每一个 S 期只能复制一次。在哺乳类动物细胞，CDK2-cyclin A 激发 G1 早期组装的复制前复合体，启动 DNA 的复制，并防止细胞重复组装新的复制前复合体。

（二）G2 到 M 期的转折使细胞进入有丝分裂期

DNA 合成完成后，细胞通过 G2 期进入有丝分裂期。有丝分裂期的 CDK 复合体是在 S 期和 G2 期形成的，但在 DNA 合成完成之前，此复合体处于无活性状态。

Cyclin B 是有丝分裂期细胞周期蛋白。在人源细胞中，cyclin B 在有丝分裂的早期核膜破裂之前进入细胞核。目前已知，促有丝分裂因子（mitosis-promoting factor，MPF）是调节所有真核细胞有丝分裂启动的关键因子，而 MPF 的本质就是 cyclin B 和 CDK1 形成的蛋白复合体。MPF 被激活后，便可引起染色质致密化，核膜破裂，有丝分裂纺锤体组装，染色体在中期板排列等分裂象改变。当所有的染色体与纺锤体微管连接完毕，MPF 又可激活细胞分裂后期促进复合体（anaphase promoting complex，APC），APC 介导分裂后期抑制因子的泛素化，进而灭活连接中期姐妹染色体的蛋白复合体。这些抑制因子的降解，使细胞周期进入到有丝分裂后期，姐妹染色体分离到相反的纺锤体极。在分裂后期的晚期，APC 促进 cyclin B 的降解，致使有丝分裂期 CDK 活性下降，进而在分裂末期分离的染色体开始松散，子代细胞核重新形成。随后，胞质分裂，产生两个子细胞。

（三）细胞周期蛋白激酶抑制因子在细胞周期调控中发挥重要作用

在 20 世纪 90 年代初，科学家们分别发现了细胞周期的重要调节因子——细胞周期蛋白激酶抑制因子，又称为 CDK 抑制蛋白（cyclin kinase inhibitor，CKI，or CDK inhibitory protein，CIP）。根据这些因子与 CDK 相互作用的特异性和序列同源性，可以分为两大家族。一类是 INK4（inhibitors of kinase 4，激酶 4 抑制因子）家族，另一类是 CIP/KIP 家族。

INK4 家族的 CKI 包括 p15（INK4b）、p16（INK4a）、p18（INK4c）和 p19（INK4d）。这些蛋白在结构上都含有 4 个锚（ankyrin）重复序列。它们可以识别 CDK4 和 CDK6，但不识别 CDK2。p16 通过与 cyclin D 竞争结合 CDK4 而引起细胞周期在 G1 期停滞。p16 的肿瘤抑制功能也是由于其可以结合 CDK4 和 CDK6，抑制 CDK-cyclin D 复合体的催化活性。p15 和 p16 经常在各种肿瘤和癌细胞系缺失，而 p18 并不具有明显的肿瘤抑制功能。

CKI 的第二个家族是 CIP/KIP，包括 p21（CIP1/WAF1/SDI）、p27（KIP1）和 p57（KIP2）。这一家族的蛋白都有一个 N 端的 CDK 抑制性功能域。它们与 CDK-cyclin 复合体相互作用，抑制 CDK2-cyclin A 和 CDK2-cyclin E 的激酶活性。CIP/KIP 抑制因子的高表达可引起 G1 期细胞周期停滞，提示它们的主要作用靶点是 G1 期 CDK-cyclin 复合体。p21 是哺乳类动物细胞中第一个被发现的 CKI 分子，其启动子上有 p53 的结合位点，受 p53 调节转录。p21 是细胞对 DNA 损伤反应时 G1 期停滞的决定因素，并且在 G2/M 期转折中起重要作用。p27 蛋白也能抑制 G1 期细胞周期蛋白-CDK 活性，引起 G1 期停滞。

近年来，一种新的 CKI 引起了人们的注意，这就是 14-3-3σ。14-3-3σ 是 14-3-3 家族成

员之一，是人类乳腺上皮细胞的特异性标记，参与 DNA 损伤后细胞周期 G2 期的调控。G2 期停滞时，14-3-3σ 使 cyclin B-CDC2 复合物隐蔽于细胞质。当 14-3-3σ 缺陷时，cyclin B-CDC2 复合体进入细胞核，引起有丝分裂的变化。这说明调节细胞周期的另一机制是改变细胞周期调节因子的亚细胞定位。

细胞周期是受到精细复杂调控的，除了转录调控使许多细胞周期调节蛋白有序地表达外，细胞周期调控还涉及两种翻译后调节机制，一是蛋白磷酸化，二是泛素依赖的蛋白降解。蛋白降解是不可逆的，保证了细胞周期过程的单向性。磷酸化则是可逆的，广泛表现于酶活性和多蛋白复合体组装的调节。蛋白磷酸化和降解又是紧密联系的，很多蛋白由于磷酸化而成为蛋白泛素化降解的靶标，同时磷酸化又调节蛋白降解机制。而且很多蛋白激酶及调节因子在细胞周期的特定阶段被降解。另一方面，蛋白激酶和磷酸化酶活性决定了在一特定时间某一蛋白的磷酸化状态。同时，磷酸酶和激酶本身也受到复杂精细的调控。此外，细胞周期调节蛋白的亚细胞定位与细胞周期调节的关系也引起人们的广泛注意。

五、生长因子具有调节细胞生长与分化的作用

生长因子（growth factors）是一类由细胞分泌的、类似于激素的信号分子，多数为肽类（含蛋白类）物质，具有调节细胞生长与分化的作用。生长因子的作用机制相当复杂，与细胞生长、分化、免疫、肿瘤、创伤愈合等多种生理及病理状态有关，因而受到科学家们极大的重视。

（一）神经生长因子是最早被发现的生长因子

1948 年，E. Bueker 等发现将小鼠肉瘤组织植入胚胎体壁可使移植区神经节增加。随后，R. Levi-Montolcini 等发现肉瘤组织的植入不仅可使局部神经节增加，而且可使远隔部位的神经节增加。由此设想肉瘤组织释放了一种可扩散因子作用于远隔部位。后来证实这种因子就是神经生长因子（nerve growth factor，NGF），它有刺激神经元生长以及神经纤维延长的功能。1959 年 S. Cohen 发现了表皮生长因子（epidermal growth factor，EGF）。R. Levi-Montolcini 和 S. Cohen 由于在这一领域的成就荣获了 1986 年诺贝尔生理学或医学奖。

目前已发现的肽类生长因子有数十种，而且还在不断增加。对大部分生长因子的结构与功能也了解得相当清楚。大多数生长因子具有促进靶细胞生长的功能，少数具有负调节功能，还有一些具有正、负双重调节作用。

（二）生长因子主要通过旁分泌或自分泌形式发挥作用

生长因子主要通过旁分泌形式作用于邻近细胞，通过自分泌作用于自身。甚至不分泌到细胞外，即可作用于自身（表 9-2）。

生长因子可来源于多种不同组织，其靶细胞亦各不相同。有的生长因子作用的细胞比较单一，如 EPO 及血管内皮生长因子（vascular endothelial growth factor，VEGF），分别作用于红细胞系和血管内皮细胞；也有的生长因子作用的细胞谱型比较广，如成纤维细胞生长因子（fibroblast growth factor，FGF）对间充质细胞、内分泌细胞和神经系统细胞都有作用。

同一生长因子对不同细胞的作用有所不同，如 HGF 对正常肝细胞的生长起促进作用，但对肝癌细胞则有抑制增殖作用。一种细胞也可受不同生长因子调节，如胚胎时属于间充质细胞的成纤维细胞可被 EGF、IGF 和多种 FGF 所调节，但不被 HGF 调节。还有一些以前认为作用比较单一的生长因子，近来发现对其他细胞也有作用。如 ET 除了对内皮细胞的作用外，可能对脑、垂体的神经内分泌也有作用。

表 9-2　主要生长因子的细胞来源和靶细胞举例

名　称	细胞来源	靶细胞
神经生长因子（nerve growth factor，NGF）	颌下腺	神经细胞
表皮生长因子（epidermal growth factor，EGF）	颌下腺	表皮细胞
转化生长因子（transforming growth factor，TGF）	肿瘤等恶性转化细胞	成纤维细胞
胰岛素样生长因子（insulin like growth factor，IGF）	肝细胞	多种细胞
肝细胞生长因子（hepatocyte growth factor，HGF）	胎盘或再生肝	肝细胞
内皮素（endothelin，ET）	血管内皮细胞	内皮细胞及血管平滑肌细胞

生长因子的生物学效应主要表现在促进细胞生长、分化、促进个体发育等方面。但是有些生长因子具有双重调节作用或负调节作用。例如，NGF 对神经系统的生长具有促进作用，但对成纤维细胞的 DNA 合成却有微弱的抑制作用。TGFβ 也是这样，对成纤维细胞有促进生长的作用，但对其他多种细胞具有抑制作用。其具体作用取决于与其他生长因子的相互作用和环境条件。

具有负调节作用的生长因子比较少，人们通常把这种负调节因子（negative growth factor）称为细胞生长抑制因子。抑素（chalone）是最早被确认的生长抑制因子。以后又发现 TGFβ、干扰素和肿瘤坏死因子等也具有抑素的某些特征，但它们实际上都是双重调节，只不过以负调节为主。目前对生长抑制因子尚无统一的标准或定义，也没有明确的学说阐明其作用机制，但是其在肿瘤、心血管等疾患的防治方面潜在的应用前景是不可否认的。因此对负调节因子的研究，始终是生物医学界的一个热点领域。

（三）生长因子通过受体发挥作用

生长因子的受体多位于靶细胞膜，为一类跨膜蛋白，多数具有蛋白激酶、特别是酪氨酸蛋白激酶的功能，也有少数为丝氨酸/苏氨酸蛋白激酶受体。最近发现细胞核也存在 EGF 等生长因子的受体样蛋白。有些生长因子受体（如 EGF 受体）与原癌基因产物有高度同源性。对生长因子受体的研究不仅有助于了解细胞的增殖分化，而且对了解肿瘤的发生、发展及治疗也具有重要意义。

1. 酪氨酸蛋白激酶受体　已知具有酪氨酸蛋白激酶功能的生长因子受体有数十种，如 EGF 受体、FGF 受体、PDGF 受体、HGF 受体（Met）、VEGF 受体等。胰岛素的受体也属于酪氨酸蛋白激酶受体。这类受体的作用方式见第八章。

2. 丝氨酸和苏氨酸蛋白激酶受体　TGF-β 家族成员通过 Ser/Thr 蛋白激酶受体转导信息。这一受体家族分成Ⅰ型受体（TGFβR-Ⅰ）和Ⅱ型受体（TGFβR-Ⅱ）。TGFβR-Ⅰ亚家族的氨基酸序列，特别在激酶结构域具有较高的相似性，而 TGFβR-Ⅱ亚家族的序列相似性则较低。TGFβR-Ⅰ和 TGFβR-Ⅱ的激酶结构域具有 Ser/The 蛋白激酶的典型序列。TGFβR-Ⅱ能进行自身磷酸化和磷酸化 TGFβR-Ⅰ的丝氨酸和苏氨酸残基，从而使受体具有了信息转导的活性。

（四）生长因子在医学实践中发挥了重要作用

经过多年研究，人们对已知生长因子的基因结构及其基因产物已经有了明确的了解。伴随着 DNA 重组技术的完善和广泛应用，以工业化规模生产原本昂贵的肽类生长因子已成为现实。不少生长因子已经应用于临床治疗，其应用前景引起广大科研和医务工作者的

高度重视。

1. 生长因子在肿瘤治疗中的作用　细胞周期异常与多种人类疾病相关，其中最重要的莫过于与肿瘤的关系。肿瘤发生的主要原因是细胞周期失调后导致的细胞无限制增殖。从分子水平来看，则是由于基因突变致使细胞周期中某些生长因子的活化或抑制，造成细胞周期调节失控。肿瘤细胞可自分泌多种生长因子，以促进自身生长。因而应用免疫学及反义核酸或 RNA 干扰等技术抑制生长因子的活性或其生成，也有可能抑制肿瘤。有些生长因子具有负调控作用，因此，直接应用某些生长因子本身就可抑制肿瘤。例如，EGF 可抑制某些 EGFR 过度表达的人乳腺癌细胞的增殖，诱导其凋亡；高浓度的 HGF 对肝癌细胞有抑制效应；NGF 可抑制神经源性肿瘤细胞的生长。

此外，某些生长因子还可以在肿瘤的诊断方面发挥作用。例如，IGF2 基因在胎肝和新生儿肝中有大量表达，但出生后该基因表达迅速降低，当肝细胞发生癌变时 IGF2 又呈胚胎性的过量表达。因此，IGF2 可以作为一个检测指标，以 RT-PCR 的方法扩增肝组织及外周血中 IGF2 mRNA 的表达水平，为肝癌诊断、转移预测和鉴别诊断等提供参考依据。

2. NGF 在神经系统疾病防治中的应用　NGF、BDGF 等能阻止神经细胞的死亡并诱导或维持其分化。NGF 在临床上主要用于外周神经损伤、视神经损伤、糖尿病患者的手足神经损伤等。它可以促进相应的神经纤维再生，有助于机体功能的恢复。另外在治疗早老性痴呆方面，可以保护中枢胆碱能神经元，防止其发生退行性病变。NGF 还可用于神经组织移植，提高植入神经元的存活率，促进它们与宿主间形成功能性联系；还可以用于脑缺血性疾病、末梢周围神经炎、自主神经功能紊乱、脑功能发育不全的治疗等。

3. 生长因子在促进创伤愈合、组织再生中的应用　有很多生长因子具有显著的促进创伤愈合的作用。如：EGF、PDGF、TGF-β 和碱性成纤维细胞生长因子（basic fibroblast growth factor，bFGF）等。其中 bFGF 已经在临床上应用得非常广泛。bFGF 是一种水溶性非糖基化蛋白质，是一种体内广泛存在、含量极微的生物活性物质。bFGF 是一种很强的丝裂原，可以促进多种细胞的增殖分化，具有促进各种软组织损伤修复的作用，可用于创伤、烧伤、溃疡等的治疗；它还可以促进毛细血管的生长和侧支循环的建立，有利于改善局部组织的微循环，参与胃溃疡的愈合过程以及参与心肌缺血与心肌梗死的修复过程；它还可以促进神经组织再生，加速神经损伤的修复。

另外，胰岛素样生长因子在促进生长发育、VEGF 等多种生长因子通过促进血管形成治疗心脑血管疾病的应用，均显示出生长因子在多种疾病防治中的应用前景。

第二节　细胞分化的分子机制

在胚胎发育过程中，一个受精卵经历多次细胞分裂不仅导致细胞数量增加，同时也诱发了各种类型细胞的出现。细胞分化（cell differentiation）便是指胚胎细胞获得不同形态、结构和功能特征的过程。人的受精卵细胞同样如此，通过限制、定向与分化等各种事件，最后产生构成人体各种组织、器官的 200 多种不同类型的细胞。细胞分化过程中的核心是基因的选择性表达，通过不同基因表达的开启或关闭，最终产生分化细胞所特有的蛋白质。

一、细胞分化是一渐进过程并受到一系列细胞信号分子的调控

细胞分化受到一系列的信号分子调控。这些信号分子通过不同的途径决定了细胞特异

性的基因表达方式。细胞分化基因的表达调控可以发生在转录、翻译以及翻译后蛋白质修饰等不同环节，其中，发生在转录水平的调节是最重要的。在通常情况下，多数与分化发育相关的基因处于抑制状态，需要在特定的发育阶段激活。基因调节蛋白或转录因子便充当了这种激活物。转录调节的一个重要特点是不同调节蛋白或转录因子之间以及它们与其他蛋白或某些小分子之间的相互作用。

（一）细胞分化贯穿于生命的全过程并伴随着特异性基因的阶段性表达

细胞分化通常是一个渐进的过程，不仅发生在胚胎时期，而且贯穿于生命的全过程。对大多数细胞来说，当其发育过程通过某一或某些关键时刻或阶段时，这个细胞的发育选择就被限定了。在卵裂阶段，内细胞团（inner cell mass）和滋养层细胞（trophectoderm）形成，随后内胚层（endoderm）、中胚层（mesoderm）和外胚层（ectoderm）形成，这些都是哺乳类动物胚胎发育过程的早期限制（restriction）事件。在三个胚层形成后，由于细胞所处的微环境的差异，细胞的分化潜能受到限制。当一个发育着的细胞通过最后一个限制点时，它的命运就被固定了。例如，某些细胞将最终转变为肌细胞或神经细胞，就是细胞的定向或细胞决定（determination）。已经被定向的细胞最终转变为成熟细胞仍需要经历一定的发育阶段。但是，通常情况下它不能从一种细胞发育轨道跳跃至另一细胞发育轨道，例如已被定向决定了的骨骼肌细胞一般不会转化为神经细胞。细胞经过定向再转化为结构与功能上特异的细胞的过程称为分化（differentiation）。近几年人们提出横向分化或转分化（transdifferentiation）的概念，即：已经被定向或已经分化了的细胞在特定的条件下可以转变为另一类功能、形态完全不同的细胞。

伴随细胞分化过程的是特异性基因的阶段性表达，例如：珠蛋白基因。成年珠蛋白有四条肽链，2 条 α 链和 2 条 β 链，即 $\alpha_2\beta_2$ 四聚体。α 与 β 珠蛋白的编码基因分别位于第 16、11 号染色体。β 珠蛋白的编码基因是一个基因家族，分别编码 β，δ，γ^A，γ^G 和 ε。在胚胎发育早期，胚胎型 ε 珠蛋白短暂性表达，随后是胎儿型 γ 珠蛋白的表达，至出生后才形成 $\alpha_2\beta_2$。有研究表明，特异性基因的阶段性表达受到这些基因在染色体上排列位置的影响。改变它们在染色体上的排列顺序，就可以丧失其发育阶段性特征。此外，特异性基因的阶段性表达还受到组合调控（combinational control）的影响，即多个基因表达调节蛋白共同协调的对某一特定基因进行调控，以及多种细胞信号分子的网络式调节。

（二）细胞信号分子在细胞分化过程中形成网络式调节

细胞分化是一个非常复杂的过程，需要多种细胞信号分子的精细协调，涉及极其复杂的信号转导过程。除了第八章中提到的信号途径以外，还有 Wnt、Hedgehog、Notch 等信号途径在细胞分化过程中发挥了非常重要的作用。而且这些信号途径之间又互相调节。

1. Wnt 信号途径　Wnt 是一类分泌型糖蛋白，通过自分泌或旁分泌等方式发挥作用。Wnt 信号途径的激活包括：Wnt 蛋白与胞膜受体卷曲蛋白（frizzled，Frz）及辅助受体（co-receptor）低密度脂蛋白受体相关蛋白-5/6（low-density lipoprotein-receptor-related protein，LRP）结合，将 Wnt 信号传入胞内，活化胞质内的 Dishevelled 蛋白（Dsh 或 Dvl），阻断糖原合酶激酶 3β（glycogen synthase kinase 3β，GSK3β）对 β-连环素（β-catenin）的降解途径，从而使 β-catenin 在细胞质中积累并发生核转位进入细胞核，与 T 细胞因子（T cell factor/lymphoid enhancer factor，TCF/LEF）相互作用，调节靶基因的表达（图 9-1）。Wnt 信号在细胞生长和分化以及多种人类疾病发生的病理过程中发挥重要作用。

2. Notch 信号途径　Notch 信号途径由 Notch 受体、Notch 配体即 DSL 蛋白（果蝇中为 Delta 和 Serrate，线虫为 Lag-2，取首字母即 DSL；在哺乳动物中为 Jagged）和 DNA 结合蛋白（在哺乳动物中为 CBF1，在果蝇中为 suppressor of hairless，在线虫中为 Lag-1，取首写字

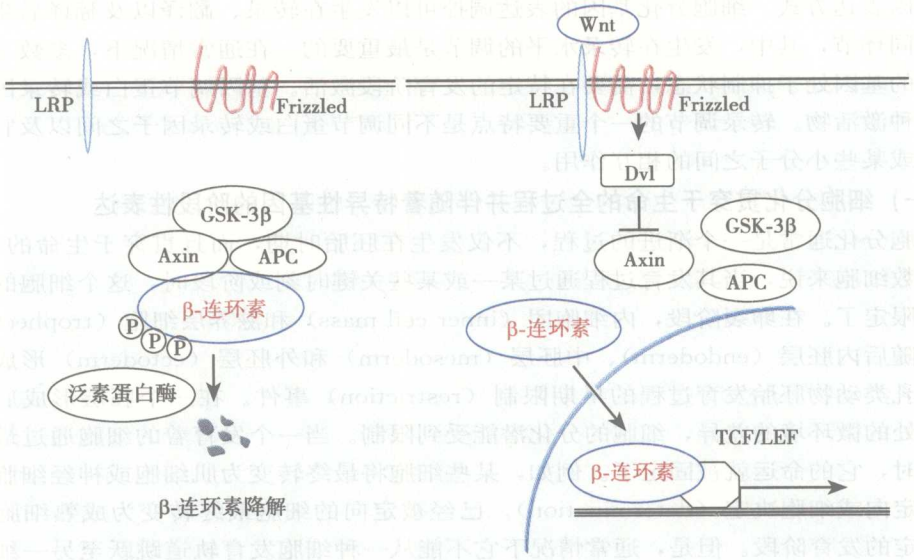


图 9-1　Wnt 信号通路示意图

左图：在 Wnt 信号分子不存在的情况下，胞内 β-连环素受 GSK-3β 等激酶磷酸化，随后被泛素蛋白酶识别，进入蛋白酶体降解。右图：当 Wnt 信号存在时，Wnt 蛋白与胞膜受体卷曲蛋白及辅助受体 LRP 结合，将 Wnt 信号传入胞内，活化胞质内的 Dvl 蛋白，阴断 β-连环素的降解途径，β-catenin 在细胞质中积累并入核，调节靶基因的表达

母即 CSL)组成。Notch 及其配体均为单次跨膜蛋白，且常常表达于同一细胞表面。当 Notch 受体和相邻细胞的配体(如 Delta)结合后，Notch 被蛋白酶两次切割，释放出具有核定位信号的胞内区(intracellular domain of Notch，ICN)，后者进入细胞核与 CLS 结合，调节靶基因表达(图 9-2)。来自相同细胞群的 Notch 信号可以产生旁抑制效应(lateral inhibition)，即当一群具有同样“命运”的细胞(equipotent cell)中有一个细胞发生了分化并表达一个特定分化信号时，这种信号抑制其周围的细胞再发生同样的分化。因此，Notch 信号在细胞的定向发育和成熟过程中发挥重要的作用。

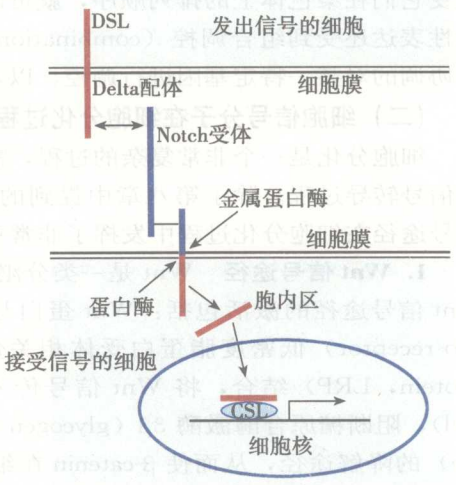


图 9-2　Notch 信号通路示意图

当 Notch 受体与相邻细胞膜上的 Notch 配体分子相结合时，细胞外基质中的金属蛋白酶对 Notch 受体进行第一次切割；随后，由胞内蛋白酶进行第二次切割，释放出具有核定位信号的胞内区，后者进入细胞核与 DNA 结合蛋白 CLS 结合，调节靶基因转录

3. Hedgehog 信号途径　Hedgehog (Hh)是一类共价结合胆固醇的分泌性蛋白，在脊椎动物中至少有 3 个基因编码，即：Shh (sonic hedgehog)、Ihh (Indian hedgehog) 和 Dhh (desert hedgehog)。

Hedgehog 信号由一个 12 次跨膜受体蛋白 Patched(Ptc)和一个七跨膜受体蛋白 Smoothened (Smo) 共同介导向胞内传递。Hedgehog 信号分子与 Ptc 结合，解除 Ptc 对 Smo 的抑制作用，进而导致下游转录因子 Gli (glioma-associated oncogene homolog) 与胞浆内的抑制蛋白解离，进入细胞核中，启动相关基因转录 (图 9-3)。

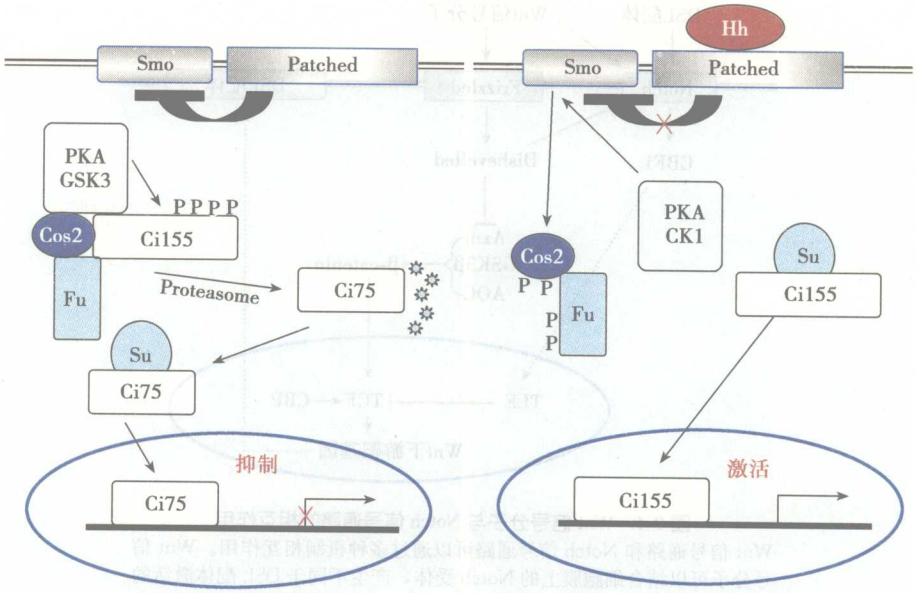


图 9-3　Hedgehog 信号通路示意图

左图：在没有 Hedgehog（Hh）信号时，Ptc 抑制 Smo，其下游转录因 Ci（Ci155）以全长形式与 Cos2、Fu 和 Su（fu）形成复合体，并募集 PKA 和 GSK3 等磷酸激酶使 Ci 磷酸化，进而经泛素/蛋白酶体途径降解，形成具有抑制活性的 Ci75。Ci75 进入细胞核后发挥抑制子的作用，抑制靶基因的转录。右图：在 Hh 信号存在的情况下，信号分子与 Ptc 结合，介导 PKA 等磷酸激酶使 Smo 磷酸化，解除 Ptc 对 Smo 的抑制作用。活化的 Smo 促进 PKA 和 GSK3 等磷酸激酶与 Cos2-Fu-Ci 复合体解离，抑制 Ci155 的磷酸化；同时，Cos2 和 Fu 被磷酸化。Ci155 进入细胞核中，启动相关基因转录

4. 不同信号通路之间的相互作用　在第八章中我们曾经简要介绍了不同信号转导通路之间的影响和调节，在细胞分化过程中不同信号通路之间的相互影响更为突出。例如，细胞外 Wnt 信号分子可能与 Notch 受体结合，从而对 Notch 通路产生不同于 DSL 配体的激活；胞质中 Dishevelled 蛋白也可以结合 Notch 受体的胞内区，抑制 Notch 通路的激活；在细胞核中 Notch 的靶基因产物 TLE 又可能通过 TCF 抑制 Wnt 信号通路（图 9-4）。

Wnt 信号通路与 Hh 信号通路之间也存在着类似的相互作用。这两条信号通路拥有一个共同的中间信号分子 GSK3，并有 G 蛋白偶联受体的参与。在上皮细胞，Wnt 信号通路中的 β-catenin 的激活可以使 Hh 信号通路的 Shh 表达上调；而 β-catenin 的抑制可以使 Shh 的表达下调。同样，Gli 对 Wnt 分子的表达也具有正调控作用。当然，在 Wnt 和 Hh 信号通路之间也存在着负调控机制。正是这些信号通路之间的精细调节才使得细胞分化有序地进行。

二、胚胎干细胞分化具有全能性

（一）胚胎干细胞具有分化为体内大多数细胞类型的能力

受精卵能够分化形成生物个体内全部的不同类型细胞的能力称为全能性（totipotency）。胚胎干细胞（embryonic stem cell）是最早期的未分化细胞，从理论上讲，它具有“全能性”。当受精卵分裂发育成囊胚时，将内细胞团分离出来进行培养，在一定条件下，这些细胞既可在体外“无限期”地增殖传代，同时还保持其全能性。1981 年，英国科学家首次成功地在体外建立了小鼠胚胎干细胞培养体系，证实了这些来自于内细胞团的胚胎

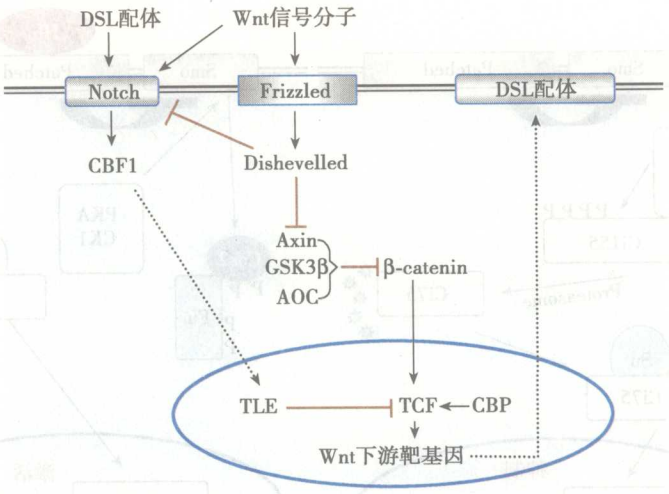


图 9-4　Wnt 信号分子与 Notch 信号通路的相互作用

Wnt 信号通路和 Notch 信号通路可以通过多种机制相互作用。Wnt 信号分子可以结合细胞膜上的 Notch 受体，产生不同于 DSL 配体激活的信号。Wnt 信号分子还可以通过 Frizzled，活化胞质内的 Dishevelled 蛋白，进而结合并抑制 Notch 活性。Notch 信号通路的靶基因产物 TLE 蛋白可以结合 TCF 抑制 Wnt 靶基因的转录。相反，Notch 配体的表达以及相邻细胞 Notch 通路的激活又可以被 Wnt 信号通路调控

干细胞可以分化为三个胚层的不同细胞。小鼠胚胎干细胞的建立为科学家提供了一个非常好的研究胚胎发育过程的模型。人的胚胎干细胞的体外培养直到 1998 年才获得成功。然而，从胚胎干细胞向不同组织细胞“定向分化”的条件至今还不清楚，胚胎干细胞分化发育的调控机制一直是干细胞研究领域的主要课题之一。

细胞发育分化的多能性（pluripotency）可定义为细胞分化为体内大多数细胞类型的能力。有的科学家认为，即便是来源于内细胞团的胚胎干细胞，也已经在某种程度上具有了一定的定向。这些细胞所具有的多能性，也足以使它们能够作为体外研究哺乳类动物发育的模型。我们知道，细胞的增殖与分化是矛盾的两个方面。处于未分化状态的细胞可以有很强的增殖能力，一旦分化为成熟细胞通常就失去了增殖能力。干细胞也是如此。在长期体外培养中如何保持胚胎干细胞处于未分化状态，并进行不断地自我增殖是一个关键的问题。

（二）维持胚胎干细胞全能性的机制非常复杂

胚胎干细胞可以进行不对称分裂（asymmetric division）和对称分裂（symmetric division）。不对称分裂是指细胞在分裂过程中，子细胞有差异的继承亲代细胞的物质和特征，一个细胞保持亲代细胞的特征，而另外一个细胞则被定向分化。对称分裂是指子细胞平等的继承亲代细胞的物质和特征，在条件合适的情况下进行快速的无限扩增。那么，是什么因素使干细胞在不对称分裂过程中出现细胞的定向？又是什么原因能够使细胞在快速增殖过程中保持未分化的表型？人们从实验中发现，含有生存必需的代谢物和营养物质的培养液并不足以维持胚胎干细胞处于未分化状态，而与滋养层（feeder）细胞的共培养是必需的，提示滋养层细胞可能分泌某种因子，维持胚胎干细胞的未分化状态。这一推测后来得到了证实：白血病抑制因子（leukemia inhibitory factor，LIF）便是发挥了这种作用。在缺乏滋养层细胞时，加入 LIF 仍然可以使胚胎干细胞保持未分化的自我更新状态。LIF 由滋养层细胞产生，缺乏 *lif* 基因的滋养层细胞不能有效地支持胚胎干细胞的增殖。撤除

LIF 滋养层细胞，胚胎干细胞继续增殖，但开始分化，并失去发育的全能性，很快死亡。

除了外部信号对维持胚胎干细胞发育的全能性或多能性具有重要作用外，细胞内部的一些决定因子的作用也是不容忽视的，尽管我们对此了解得还不十分清楚。目前，只有 POU 家族转录因子 Oct-3（也称为 Oct-4，Oct-3/4 或 NF-A3，由 *pou5f1* 编码）在维持细胞多能性中的作用得到了肯定。

Oct-3/4 最初是在胚胎瘤细胞中发现的。Oct-3/4 的表达严格局限于小鼠的全能或多能细胞中，如卵母细胞、早期分裂期胚胎、囊胚的内细胞团、卵圆柱期胚胎的原始外胚层（primitive ectoderm，PEC）及原始生殖细胞（primordial germ cell，PGC），提示它在保持细胞发育多能性中的作用。Oct-3/4 的表达水平决定了胚胎干细胞三种不同的命运。在胚胎干细胞内，Oct-3/4 的水平只有维持在特定水平，才保持它能处于不断自我更新状态；而当 Oct-3/4 水平是正常的 2 倍时，胚胎干细胞分化为原始内胚层和中胚层细胞；当 Oct-3/4 水平减少到正常的 50%时，胚胎干细胞则去分化为滋养层细胞。在 Oct-3/4 基因敲除的小鼠，桑椹胚的细胞分化为滋养外胚层。

Oct-3/4 的表达调控是一个复杂的机制，有多个信号分子参与。2006 年，新加坡国立大学的 Y. Loh 等人报导 Oct-4 和另外一个转录因子 Nanog 通过一种网络式的调节维持胚胎干细胞的全能性，并利用 chromatin immunoprecipitation paired-end ditags（PET）的方法在小鼠的基因组中发现了 Oct4 的 1 083 个结合位点，Nanog 的 3 006 结合位点。然而 Oct4 和 Nanog 如何调控如此众多的基因来共同维持干细胞的全能性和增殖能力还并不清楚。这个领域的研究还刚刚起步，我们相信这一领域的研究将有助于全面了解胚胎干细胞自我增殖及定向分化的机制。

三、成体细胞的横向分化

横向分化又称为转分化（transdifferentiation）。通常是指一种已经分化的细胞类型不可逆地转化为另外一种正常分化的细胞类型。转分化前后的细胞在形态上和分子水平上都有明确的特征以示区别。在一些情况下，转分化伴有细胞分裂，而在另外一些情况下，不伴有细胞分裂。转分化不可避免地伴有不同基因的关闭或启动，也就是基因组活动的重新编程（reprogramming）。我们这里所讨论的转分化是指成体干细胞或前体细胞/祖细胞的进一步分化。

成体干细胞（adult stem cell）是指存在于已经分化的组织中但自身尚未完全分化，且具有自我更新、在特定条件下能够进一步分化的细胞。人体几乎所有组织都存在成体干细胞。成体干细胞已经有相当程度的分化，如果不受外加条件的影响，一种组织的成体干细胞倾向于分化成该组织的各种细胞，比如造血干细胞在体内自动分化成各种血细胞。但是，近几年的研究结果证实：在特定的条件下，一种组织的成体干细胞可以转分化成其他组织的功能细胞。例如造血干细胞可以分化成神经细胞和肝细胞；间充质干细胞可以分化成神经、肌、软骨和骨等多种细胞；脂肪组织的干细胞，在一定条件下可以分化成神经细胞。转分化现象的发现促使人们重新认识细胞发育分化的机制，也使得利用病人自身健康组织的干细胞来修复病损组织成为可能。不过，也有不支持成体干细胞转分化的观点。他们认为，人们所观察到的转分化现象可能是细胞融合（cell fusion）的结果。

细胞的终末分化状态往往被认为是固定的。而细胞的这种特定状态是由一些转录因子的共同作用决定的。理论上，通过改变转录因子的表达，任何一种细胞都有可能被变为另外一种细胞，而在某些情况下，决定的状态是由一种主要的因子所控制。例如，MyoD 家族的 bHLH 因子，能够迫使培养中的许多细胞类型向肌细胞分化。编码这些因子的基因

往往被称为主控制基因（master control gene）。事实上它们的作用在一定程度上还要依赖于其他转录因子的活性。

2006年，日本科学家发现，用4个转录因子（Oct3/4、Sox2、c-Myc、Klf4）过量表达可以将小鼠成纤维细胞逆转到细胞分化前的状态，从而获得功能上与胚胎干细胞类似的诱导性多能干细胞（induced pluripotent stem cell，iPS）。2007年，这个小组的科学家用同样的方法成功的获得人的iPS。与此同时，美国科学家利用Oct4、Sox2、Nanog、Lin28四种转录因子也从人的成纤维细胞中诱导获得了iPS。这些研究表明，某些主控制基因可以决定细胞的分化或未分化状态。

四、再生医学是21世纪医学研究的热点

组织再生是生物界普遍存在的现象。不同种属动物的再生潜能差异很大，与无脊椎动物相比，脊椎动物特别是高等哺乳动物的再生能力非常有限。此外，同一种属动物不同组织的再生能力也同样具有相当大的差异。例如，人的上皮组织、骨组织和肝组织等具有较强的再生能力；而神经组织、心肌组织等则一直被认为是不能再生的组织。尽管近年来有些研究报道神经组织和心肌组织又具备一定的再生能力，但是这些组织的再生能力很弱是一个不争的事实。因此，具有多种分化潜能的干细胞便引起人们广泛的关注。人们期待着利用干细胞的这种特性进行组织再生，实现对机体组织损伤的修复，再生医学（regenerative medicine）应运而生。

再生医学是一门通过研究机体正常组织特征与功能、损伤修复与再生机制，进而寻找有效的生物治疗方法，促进机体自我修复与再生，或构建新的组织与器官，以改善或恢复损伤组织和器官之功能的科学。再生医学是一门新兴的多学科交叉学科，涉及医学与生命科学、材料科学、力学、数学、计算机科学和工程学等不同学科。利用这些学科的原理与方法，人们已经研发出各种产品，用于修复、改善或替代人体组织器官的产品，并应用于临床治疗某些因疾病、创伤所造成的组织器官缺损或功能障碍，并显示出良好的发展前景。

（周春燕）

第二篇　疾病的分子生物学基础

人体各种疾病的发生都有其相应的分子基础。从理论上讲，任何疾病都与一种或多种基因的基因型或其表型直接或间接相关。疾病与基因型相关的含义是疾病的发生和发展直接与基因或基因组中存在的变异或多态性有关。疾病与基因表型相关的含义一方面是基因型的变异在表达产物的体现，另一方面，环境中的刺激首先改变细胞内蛋白质修饰状态，进而改变细胞的功能状态而导致疾病，这种疾病常常涉及细胞内信号转导分子的异常。在表型方面的异常仍然会作用在基因型方面，导致细胞基因型的变异。

本篇将从疾病的遗传机制出发，分别讨论细胞生长异常的分子机制、病原微生物致病、炎症、应激反应和衰老的分子机制，并介绍心血管系统、免疫系统和内分泌系统部分疾病的分子机制。需要指出的是，上述分子机制大都在近年来才有所认识，而且目前的认识尚不够系统和全面，故在介绍中难免以偏概全。有些异常的分子机制可能目前仅涉及个例，而非普遍现象，不过考虑到这些问题往往对于各系统疾病分子机制的认识会给予一定带有启发性的思考，故仍然占有少量篇幅。

部分章节在讨论疾病的分子异常之前，介绍了该系统的发育分子生物学进展。这些有关细胞分化和器官发生的进展虽非讨论疾病本身，但是对于认识疾病的分子异常往往具有一定提示作用，尤其是对许多先天性疾病的发生机制的认识将有极大的帮助。此外，由于干细胞在疾病治疗中的不可估量的前景，对细胞分化机制的理解对于医学生的重要性也是毋庸置疑的。

本篇侧重讨论的是疾病的分子机制，即从分子生物学角度分析疾病发生、发展及转归，并不涉及疾病的临床表现、诊断和治疗方案。

第十章　基因变异与疾病

关键知识点

- **基因变异的基本概念**：基因变异指 DNA 的序列组成或结构在传递过程发生了改变，也称基因突变。就其在人群中的分布和生物学效应而言，可将基因变异划分为 DNA 多态性和致病突变。根据 DNA 序列变异的物理形态及其对基因功能的影响，人类的基因变异可大致划分为点突变、大片段突变、短串联重复突变和拷贝数变异四大类。每一类基因变异均包含了 DNA 多态性和致病突变。
- **基因变异的生物学效应**：一般认为，发生在编码基因的变异会引起更显著的性状变化，人类绝大

多数单基因遗传病的产生都是由于突变使受累基因的功能受损所致。就突变对基因功能影响的结果而言，可大致将基因突变划分为功能失去性突变和功能获得性突变两类，无论是基因功能的丧失/减弱或获得/增强，对于维持细胞的正常代谢和生命活动都是不利的，这是人类遗传病发生的分子基础。

- **基因突变的致病机制**：导致基因产生不同生物学效应的机制各不相同，这与基因表达的组织细胞，突变类型，突变累及的基因及其发生的位点有关。从基因表达的分子生物学基本概念上讲，无论是何种致病机制或细胞内异常代谢通路，基因突变对于基因功能的影响，其本质性生物学效应可以理解为基因表达水平的改变（下调或上调）和蛋白质结构的改变。

关键词：遗传变异；DNA 多态性；功能获得性突变；功能失去性突变

遗传和变异是通过具体性状而被认识的。性状（character，trait）是指生物所具有的形态、功能或生化的特点。生物世代相传的不是性状，而是决定性状的基因。基因是遗传的物质基础和功能单位，它包含在产生多肽链的 DNA 序列中，DNA 序列的保守性非常重要，如果在传递过程中基因的组成或结构发生改变，将引起生物性状的改变，这就是变异（variation）。基因变异是生物界普遍存在的现象，广泛分布于人体基因组中的主要有两种来源的变异：单核苷酸多态性（single nucleotide polymorphisms，SNP）和拷贝数变异（copy number variant，CNV）或称拷贝数多态性（copy-number polymorphisms，CNP）。这些基因变异不仅是人类进化的源泉，也是人类表型多样性和疾病易感性（disease susceptibility）的遗传学基础，人种间及人群中个体间不同的习性、外貌特征乃至患病的差异性等，可以理解为由这些基因变异所贡献的遗传多态性（genetic polymorphism）。限于篇幅，遗传多态性与人类多基因常见病的疾病易感性关系的内容，本章不作详细介绍。

基因变异又是人类遗传病产生的根源，导致疾病发生的基因变异称为致病突变（causative mutation）。一般认为，发生在编码基因的变异会引起更显著的性状变化，人类绝大多数单基因遗传病的产生都是由于突变使受累基因的功能受损所致。突变基因可通过生殖细胞产生的配子传递给后代，使后代产生遗传病，这就是人类遗传病的本质之一，即生殖系突变（germline mutation）。与之相对应的发生在生殖细胞以外的突变称之为体细胞突变（somatic mutation），这类突变可成为引发肿瘤的重要因素。人类基因组的致病突变可表现为 DNA 碱基的替换、缺失、插入、重排等基因结构的改变，也可以动态突变的形式出现，CNVs 也是致病的原因之一，其结果是使不能产生基因表达产物、产物的表达量不足或过量，或基因表达的蛋白质产物功能失活或激活。本章将围绕人类遗传性疾病的分子机制的主题，重点阐述基因变异的生物学效应及其与人类遗传病发生的关系。

第一节　基因变异的基本概念

基因变异是指 DNA 的序列组成或结构在传递过程发生了改变，也称基因突变。

一、基因变异包括 DNA 多态性和致病突变

依据基因变异在人群中的分布和生物学效应，我们可将其划分为 DNA 多态性和致病突变。人基因组中最普遍的 SNP 和拷贝数多态性在人群中的频率被定义为大于 1%，致病

突变的频率通常很低，多态性这一名词着重说明 DNA 序列在人群中不同个体、不同等位基因间的差异性，而致病突变则强调遗传变异会引起基因功能的显著改变。通常基因变异包含了产生生物学效应的意义，故不能将 DNA 多态性理解为非致病突变，当某些特定的 DNA 多态性被认为与疾病发生相关时，也可用结构畸变或致病突变加以描述。致病突变是用于描述那些导致人类遗传病（主要是单基因遗传病）发生的变异事件的名词，这类突变主要发生在人类基因组的结构基因上。发生在结构基因上的基因突变若不影响基因功能，称之为中性突变（neutral mutation）。

二、人基因组中有多种类型的基因变异

根据 DNA 序列变异的物理形态及其对基因功能的影响，人类的基因变异可大致划分为点突变（point mutation）、大片段突变（gross mutation）、短串联重复突变（short tandem repeat mutation，STR mutation）和拷贝数变异（CNV）四大类。每一类基因变异均包含了 DNA 多态性和致病突变，广泛存在于人基因组中的三类代表性 DNA 多态性见表 10-1。致病突变通常发生在人类的结构基因区域，并影响基因的编码功能、或使天然的基因缺如或结构被破坏而丧失功能。表 10-2 小结了一些有代表性的基因突变的常用名词及其所产生的生物学效应。

表 10-1　人基因组中三类常见的 DNA 多态性

变异类型	定义	人基因组中的频率
SNP	给定人群中基因组的给定位点发现的单个核苷酸变异，人群中出现的频率>1%	人基因组中约有 1 000 万个
STR‡	可变数目的 1～6 bp 核心单元组成的重复序列，一般总长度<200 bp	人基因组中至少有 100 万个，约占人 DNA 序列的 3%
CNV*	人类基因组中大小从≥1 kbp 到超过 Mb 不等的 DNA 片段的拷贝数变异，它包括倒位、缺失、插入重复等的拷贝数变异，人群中出现的频率>1%时，称为拷贝数多态性	目前尚不清楚，估计≥50 kbp 的 LCV 约占人 DNA 序列的 12%

‡不包括 7～100 bp 核心单元组成的可变数目串联重复序列（variable number of tandem repeats，VNTR）；* 包括长度>1 kbp 的插入缺失（InDel）；DNA 片段≥50 kbp 的称为大尺度 CNV（large-scale CNV，LCV）

拷贝数变异（CNV）是近年来在人基因组中发现的一类新的基因变异（图 10-1），研究者已经发现 CNV 在不同的人群中和同一人群的个体间较 SNP 存在着更大的差异性（可高达 12%），且已积累了一些 CNV 作为人类遗传病原因的案例，这些研究结果提示，通过系统阐述 CNV 区域所含基因及其数目对基因表达的影响，以及研究其位置或结构域效应对基因功能的影响，将为揭示人类疾病的病因学打开一扇新的窗口。

表 10-2　几种常见的结构基因突变

突变类型		定　义	生物学效应
点突变	同义突变	突变发生在三联密码子的第三位碱基	仍编码相同的氨基酸，不影响蛋白质结构，可不产生表型变化
	错义突变	突变通常涉及三联密码的头二位碱基中的一个	氨基酸发生改变，当影响蛋白质活性中心时，影响其功能，有时不影响蛋白质的活性，无明显表型改变

续表

突变类型		定义	生物学效应
点突变	无义突变	突变使编码氨基酸的密码子变成三个人类基因终止密码子之一	多肽链合成提前终止，通常产生无活性的多肽链
	剪接突变	位于内含子与外显子内或交界处影响 RNA 剪接的突变	使剪接效率降低或产生异常剪接子，抑制正常基因表达
	启动子突变	位于启动子区影响基因转录的突变	使受累基因转录水平抑制或增强
	移码突变	在 DNA 分子中插入、缺失少数几个非三的整倍数的碱基	突变位点以下的所有氨基酸序列发生改变，由于其下游出现新终止密码使产生缩短或延长的多肽链，蛋白质失活或功能异常
大片段突变	缺失	在特定位点减少了一段数十到数千个碱基对 DNA 片段	结构基因序列丢失使基因产物缺如；重排造成编码产物异常或影响基因表达
	插入	在特定位点增加了一段数十到数千个碱基对 DNA 片段	使插入点序列重排导致受累基因表达失活或激活
	重排	特定位点的 DNA 片段的排列顺序改变，如倒位（inversion）、重复（duplication）等	通常可使基因表达产物剂量增加或影响基因表达，导致细胞功能的紊乱
STR* 突变	三联体重复扩增	DNA 中的三核苷酸（如 CGG、CCG、CAG 等）重复序列的拷贝数异常扩增（2 至几千个拷贝数不等）	扩增序列一般发生在结构基因区，通过影响基因的表达导致受累基因功能异常

* 通常指 1-6 bp 核苷酸拷贝数变异（见表 10-1），三核苷酸重复序列扩增（triplet-repeat expansion）是人类遗传病中最常见的 STR 突变类型，又称动态突变，重复序列的拷贝数可随着世代的传递而不断积累而逐渐增多，这一机制可解释遗传早现（anticipation）现象和临床表现的严重程度

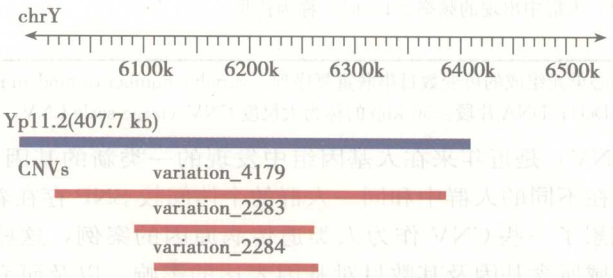


图 10-1　Y 染色体上三个 CNV 的物理图谱

三、基因变异是可以遗传的

无论是致病突变还是 DNA 多态性，只要它们是生殖系突变，即可通过生殖细胞产生的配子传递给后代，其遗传方式都遵循孟德尔定律，子代所拥有的基因变异来自于亲代的垂直传递（图 10-2）。基因变异的可遗传性是目前遗传病诊断、亲子鉴定、基于家系的连锁分析和单倍型构建的理论基础。在鉴定人类遗传病的 DNA 变异时，还有一种非遗传性的新发突变（de novo mutation），由于此类基因变异属体细胞突变，故不会传递给下一代。

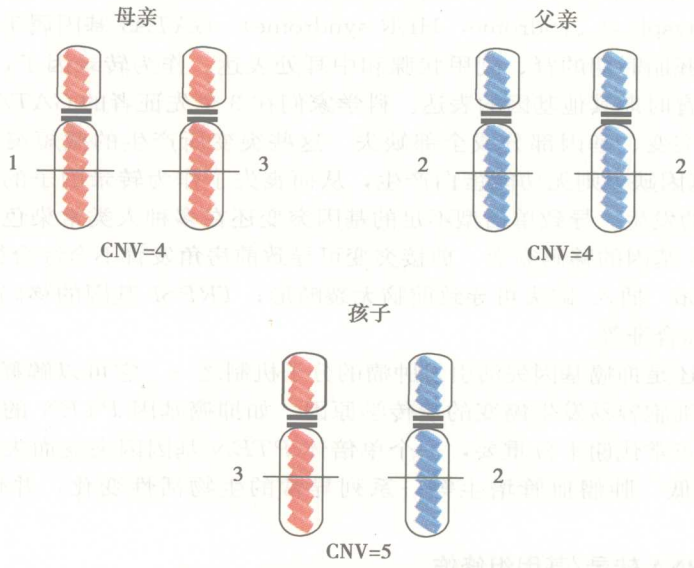


图 10-2　基因变异的垂直传递模式图

图中 CNV 示个体的变异体拷贝数。该数字也可理解为 STR 核心序列拷贝数，或不同的碱基变异

第二节　基因变异的生物学效应

就突变对基因功能的影响而言，可大致将基因突变划分为功能失去性突变（loss-of-function mutation），或功能获得性突变（gain-of-function mutation）。无论是基因功能的丧失/减弱或获得/增强，对于维持细胞的正常代谢和生命活动都是不利的，严重时就会导致疾病。基因突变导致基因产生不同生物学效应的机制各不相同，这与突变类型，突变累及的基因及其发生的部位有关，本节选取一些有代表性的导致人类遗传病发生的分子机制，来阐述突变引起基因功能减弱或增强这两类不同生物学效应的基本原理。

一、DNA 变异可使基因功能减弱

（一）单倍型不足

单倍型的概念最初源于单倍体，一个配子中全部染色体的总和可称作单倍体，人体二倍体细胞中成对染色体中的一条染色体的全部基因组可称作单倍型。分别来自父亲和母亲的两个单倍型则组成个体的基因型（genotype）。一条染色体上距离很近的一组 SNP 位点很少由于发生重组被分离而倾向于以一个整体遗传给后代。我们将一条染色体上紧密连锁的一组 SNP 位点组成的基因单位称作单倍型（haplotype）。单倍型不足（haplotype insufficiency）是指给定基因的两个拷贝中的一个等位基因发生突变或缺失，另一个拷贝的表达产物不足以维持正常的细胞功能。这一概念着重于基因表达的剂量，简单地说就是正常基因所产生的蛋白质数量不能满足细胞代谢的需要。单倍型不足是解释基因突变导致人类常染色体显性遗传病的重要机制之一。

2000 年，研究人员发现 *GATA3* 基因的单倍型不足会导致人类患上 HDR 综合征（甲

状旁腺功能减退-感觉神经性耳聋-肾发育异常综合征，hypopara-thyroidism，sensorineural deafness，renal dysplasia syndrome，HDR syndrome）。*GATA3* 基因属于锌指蛋白转录因子家族，在人类胚胎时期的肾、副甲状腺和中耳处表达。作为转录因子，*GATA3* 蛋白还参与调控胚胎发育时期其他基因的表达。科学家们在 3 个先证者的 *GATA3* 基因中发现了无义突变、移码突变、基因部分或全部缺失。这些突变所产生的截短突变型蛋白不能与 DNA 结合，而基因缺失则无功能蛋白产生，从而丧失了作为转录因子的调控功能，并导致 HDR 综合征的发生。导致单倍型不足的基因突变还在多种人类常染色体显性遗传病中发现，如 *PITX2* 基因的移码突变、剪接突变可导致前房角发育不全综合征；*ZIC2* 基因的小片段（7～56 bp）插入/缺失可导致前脑无裂畸形；*TRPS1* 基因的移码突变等可导致毛发-鼻-指（趾）综合征等。

单倍型不足还是抑癌基因失活引发肿瘤的分子机制之一。它可以解释当抑癌基因的活性剂量不足时，细胞容易发生癌变的遗传学原因。如抑癌基因 *PTEN* 的表达水平对于维持前列腺细胞的正常代谢十分重要，一个单倍体 *PTEN* 基因因突变而失活会引发包括细胞增殖、凋亡降低、肿瘤血管增生等一系列异常的生物活性变化，并促进前列腺癌的恶化。

（二）反义 RNA 转录/基因组修饰

反义 RNA 转录并介导基因组修饰是导致人类遗传病发生的一种新的分子机制，这一机制是在一个有着特殊表型的 α-地中海贫血（α-地贫）家系的研究中被阐明的。正常人一条染色体上有二个功能性 α 珠蛋白基因（*HBA2* 和 *HBA1*），该家系病例成员中均被鉴定缺失一个 *HBA1*，且 *HBA2* 基因排除了点突变对基因表达的影响，但却表现出与缺失 *HBA2* 和 *HBA1* 两个基因的典型 α-地贫杂合子病例类似的血液学特征。导致结构完整的 *HBA2* 基因失活的分子机制见图 10-3。*HBA2* 基因下游 23 kbpDNA 片段的缺失（其中包含了 *HBA1* 基因、*HBQ1* 基因和 *LUC7L* 基因的 3′ 末端最后三个外显子和多聚腺苷酸剪切位点），使其下游与 *HBA2* 基因转录方向相反的 *LUC7L* 基因靠近了 *HBA2* 基因，*LUC7L* 基因的转录由于其终止密码丢失而延伸覆盖了 *HBA2* 基因及其启动子，其反义 RNA 与 *HBA2* 的转录本形成部分双链 RNA，使 *HBA2* 的 RNA 降解，同时介导 *HBA2* 上游位点的 CpG 岛发生甲基化，这样，结构完整的 *HBA2* 基因由于反义 RNA 封闭和甲基化作用而转录沉默，α 珠蛋白链缺乏，导致 α-地贫发生。

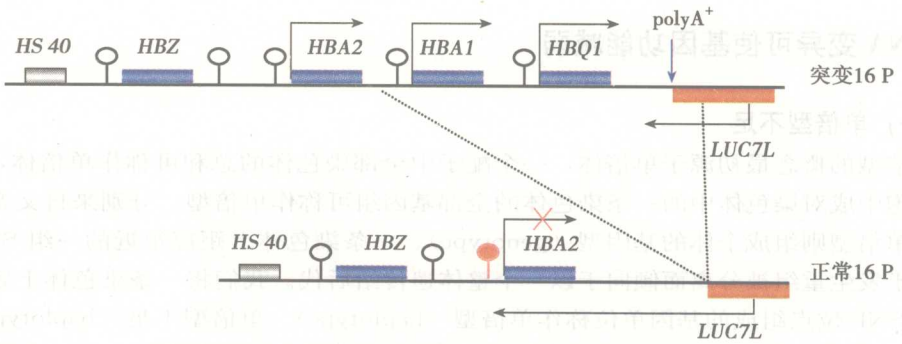


图 10-3　*HBA2* 基因失活的反义 RNA 转录分子机制

图中蓝色和红色方框示功能性结构基因（名称标在方框上或方框下），其转录方向用箭头表示；空心小圆示 CpG 岛未甲基化，红色实心小圆示 CpG 岛甲基化

（三）转录因子基因变异

转录因子一般具有特定的模体，可以与 DNA 分子结合，从而起始、促进或抑制转录

的进程。如果转录调节因子的基因本身发生突变，产生异常的调节蛋白，就会影响其调节的下游基因的正常转录，产生基因功能减弱的生物学效应。

β-珠蛋白基因的调节基因之一——*GATA-1* 基因突变所导致的 β-地中海贫血（β-地贫）是说明转录调节因子基因变异导致人类遗传病发生的良好案例。*GATA-1* 基因是一种锌指结构的转录因子，该蛋白氨基端的锌指基序负责与 DNA 分子的结合。它是造血作用中重要的调节蛋白。当 *GATA-1* 蛋白的 N 端锌指区的第 216 位氨基酸由精氨酸突变为谷氨酰胺时，蛋白与 DNA 分子的结合稳定性下降。与野生型蛋白相比，突变型蛋白虽然可以与 DNA 分子结合，但它从 DNA 分子上的解离速度快于野生型蛋白。人类 β-珠蛋白基因上游启动子区存在多个 *GATA-1* 蛋白的结合位点，证明这些基因的转录直接受到 *GATA-1* 的调控。所以，该基因的突变会造成人 β-珠蛋白基因转录水平下降，使 β-珠蛋白表达量减少而不能组成足够的血红蛋白，因而引起 β-地贫发生。转录因子基因变异导致遗传病发生的分子机制见图 10-4。

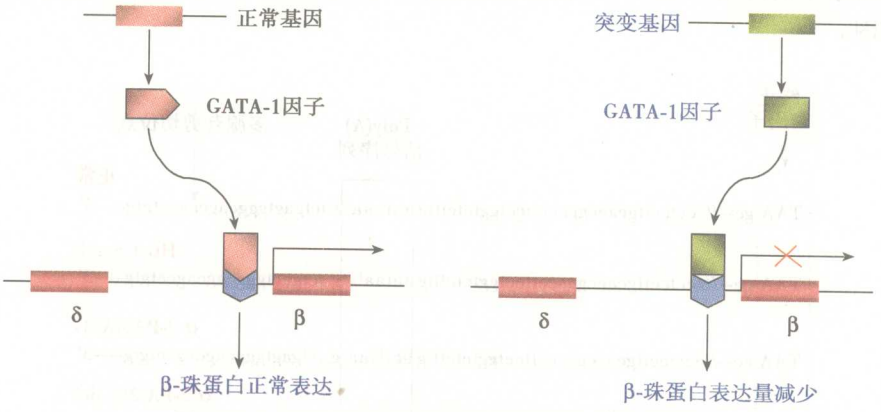


图 10-4　*GATA-1* 基因突变导致 β-地中海贫血的分子机制

图中红色方框，其中野生型 *GATA-1* 基因用红色方框表示，突变基因用绿色方框表示，β 珠蛋白基因抑制作用用红色×表示

我国科学家还发现，常见的致盲性眼病-儿童遗传性白内障的致病“祸首”也是一个转录因子的突变。人眼瞳孔后方有一透明而有弹性的凸透镜样结构，叫做晶状体，如同照相机的镜头。当它由透明变为混浊，就称为白内障，三分之一的儿童白内障属遗传性疾病。研究人员在 3 个中国人家系和一个丹麦人家系中发现，*HSF4* 基因突变是遗传性儿童白内障的致病原因。*HSF4* 基因产物是热休克蛋白（heat shock protein，Hsp）转录因子，可以调控 *HSP70*，*HSP90a* 和 *HSP27* 基因的表达，而上述 *HSP* 基因家族的产物作为分子伴侣在人类晶状体的发育和维护中发挥重要作用。*HSF4* 基因的错义突变会导致 HSF4 蛋白的 DNA 结合基序发生变化，使得该蛋白与 DNA 分子的结合能力降低，从而下调相关 *HSP* 基因的转录水平，而 Hsp 蛋白产量的下降会引发晶状体蛋白质合成、装配、降解等生理过程的障碍，最终导致遗传性儿童白内障发病。

（四）影响 mRNA 稳定性的变异

真核生物编码蛋白基因的 3′ 端非翻译区（3′-UTR）内存在一些参与 mRNA 加工的保守序列，如见于人 α-珠蛋白基因序列中的多聚腺苷酸信号序列 AAUAAA，及其下游 16 个核苷酸处的多聚腺苷酸剪切位点，在紧邻 3′-UTR 的 3′ 端旁侧区还有一些其他的 mRNA加工顺式元件，如一段 16 个碱基的 AU 富集区（AU-rich elements，AREs），这些元件与细胞核中的特异性反式作用因子结合是实现 mRNA 加工及保持其稳定性的分

子基础。此外，3′-UTR 及其旁侧序列还与其他蛋白复合体相互作用参与核转运和翻译调控等过程。

在这一保守序列区发现的一些天然突变是导致 α-地贫的原因。见于 AAUAAA 序列中的 A→G 突变（AAUAAA→AAUAAG）是世界上最先阐明的导致人类遗传病的 3′-UTR 功能减弱性突变，该突变不能完全抑制在正常多聚腺苷酸剪切位点进行 mRNA 加尾，但同时产生大量的超越了剪切位点的“加长型”mRNA 加工产物，这种异常 mRNA 会被认为是“异己”而遭到细胞内 mRNA 降解机制的迅速清除。此外，使 mRNA 加工产物不稳定的因素还有 3′-UTR 参与翻译调控的异常机制。一个经典的例子是引起 α-地贫的一种常见的异常血红蛋白（Hb Constant Spring），该变异破坏了正常的终止密码子，而由谷氨酰胺密码子所替代（UAA→CAA），肽链合成至其下游另一个终止密码子（图 10-5）处终止。其结果是，迫使核糖体进入 3′-UTR，翻译产生一个增加了 31 个氨基酸的异常 α-链。核糖体的进入使原本调节加工 mRNA 产物稳定性的 α 蛋白复合体的位置被挤占，从而加速了异常转录本的降解。mRNA 的快速降解使 mRNA 加工产物不稳定是造成其功能丧失的主要原因。

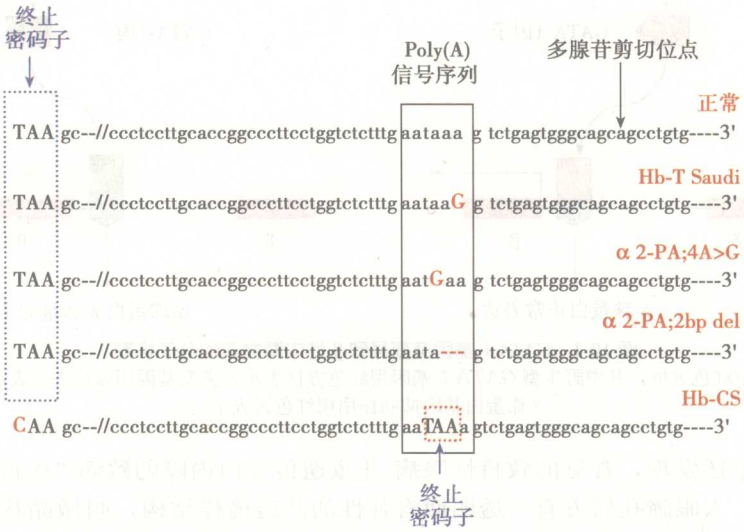


图 10-5　*HBA2* 的 3′UTR 区变异抑制转录加工或转运

（五）显性负效应

蛋白质复合体（protein complex）是细胞活动的重要功能单元，显性负效应（dominant negative effect）又称反显性效应，是指由于基因变异导致参与组成蛋白质复合体的某一成员蛋白产生了功能缺陷，该突变型蛋白虽仍能与野生型蛋白形成多蛋白质复合体，但却使该蛋白质复合体完全或部分丧失了功能，因此可引发显性表型。简单地说就是突变型蛋白通过蛋白-蛋白相互作用干扰了正常蛋白质的活性，使之形成的蛋白质复合体功能丧失或减弱。导致显性负效应的基因突变称为显性负效应突变（dominant-negative mutation），此基因编码的突变型蛋白称为显性负效应蛋白（dominant-negative protein）。显性负效应是相对显性正效应（dominant positive effect）而言的，显性正效应通常是指功能获得性突变效应，即突变型的蛋白质结构或表达方式发生改变，使该蛋白质具有新的功能，且该功能超越了其野生型等位基因所编码的蛋白质的活性，表现为显性。

显性负效应是解释部分常染色体显性遗传性长 Q-T 间期综合征的分子机制。遗传性长 Q-T 间期综合征是一种先天性发作性心脏疾病，以心电图 QT 间期异常延长（QTc＞440 ms）为主要特征，常突发严重心动过速、室颤等，并可导致晕厥甚至猝死。该病是由于心肌细胞的离子通道异常引发细胞异常电活动所致。目前至少有 7 个离子通道蛋白编码基因被鉴定为致病基因，其中人群中最常见的是 *KvLQT1*（*KCNQ1*）和 *HERG* 基因，受累基因的错义突变是导致疾病发生的主要原因，其他突变型包括移码突变、整码缺失、剪接突变和无义突变等。*KvLQT1* 基因定位于 11 p15.5，研究者通过对该基因三种天然突变型蛋白（A177 P、L272 F 和 T311 I）的功能研究，阐明了突变型蛋白通过蛋白-蛋白相互作用抑制野生型蛋白质复合体功能，进而产生显性负效应的分子机制。主导心肌电活动的膜通道蛋白的功能障碍是长 Q-T 间期综合征的分子基础。

此外，显性负效应还是解释抑癌基因变异相关的恶性肿瘤发生的重要机制。在野生型等位基因拷贝存在并表达的情况下，抑癌基因编码的突变体蛋白通过抑制野生型蛋白的活性而产生显性负效应。例如，*P53* 突变等位基因编码的 P53 蛋白突变体可以与野生型等位基因编码的 p53 蛋白形成寡聚蛋白质复合体结果导致野生型 P53 蛋白功能丧失，从而引起细胞的转化和癌变。

二、DNA 变异可使基因功能增强

（一）转录增强作用

与许多转录抑制性突变相反，一部分调节序列的变异也会增强基因转录，使基因表达水平提高，产生异常表型。如在人类遗传性胎儿血红蛋白持续增多症（hereditary persistence of fetal hemoglobin，HPFH）个体中，已经鉴定了若干种位于 γ 基因启动子区的可促进该基因转录的点突变，这些突变包括^{G}γ 启动子区－202 C→G、－175 T→C 和－114 C→T 突变，以及发生在^{A}γ 启动子区的－196 C→T、－175 T→C 和－117 G→A 等突变，这些启动子变异可上调 γ 基因的表达水平，使成人期本已关闭的表达胎儿珠蛋白链的 γ 基因重新开放，引起成人红细胞中 γ 链持续高表达状态，产生 HPFH。其中^{G}γ-202 C→G、^{A}γ－198 T→C、^{G}γ－175 T→C、^{A}γ－175 T→C、^{A}γ－117 G→A 和^{A}γ－196 C→T等有更高水平的 γ 链表达，属 Hb F 高表达的 HPFH，其他则为 γ 链表达上调幅度不大的 Hb F 低表达的 HPFH。γ 基因表达增强是由于这些突变改变了转录调节蛋白在启动子区的结合效应所致。^{A}γ－196（C→T）突变是见于中国人的一种 γ-基因启动子区突变。此外，^{G}γ-基因启动子区的多态性位点^{G}γ－158 C→T 也被鉴定与低 Hb F 的 HPFH 有关。

另一方面，一些转录因子与某些 DNA 特异性序列进行结合，可抑制基因转录活性。如果该转录因子编码基因发生变异或者与之结合的 DNA 特异序列产生突变，都会影响其结合性能，从而使该调节基因对其调节的下游靶基因失去转录抑制作用，从而产生转录增强的生物学效应。如在散发性或遗传性 Wilms 肿瘤的发生机制研究中，发现与编码转录因子的 *WT1* 基因突变有关，野生型 *WT1* 基因的表达产物是含锌指结构的蛋白质，可识别并特异性地结合在被调节基因的一段特异性碱基序列（5′-CGCCCCGC-3′），起到阻抑该基因转录的作用。而当细胞中的 *WT1* 基因发生了基因突变（如错义突变），所产生的突变型锌指蛋白就不能与该段特异性碱基序列结合，也就失去了转录调节作用，导致其下游多个靶基因，如血小板衍生生长因子 A 链基因（*PDGF-A*）、胰岛素样生长因子Ⅱ基因（*IGF-2*）、集落刺激因子 1（*CSF-1*）、维甲酸受体 α 以及 *WT1* 基因本身的转录被抑制或激活，成为引起细胞癌变的原因之一。

（二）增强子的位置效应

基因表达的调控是一个很复杂的过程，除结构基因自身的顺式调节元件和相关的反式作用因子参与外，还有一些远离结构基因的其他调节元件及其相关的细胞调节因子，如远离基因转录起始点的增强子的调节作用。增强子是存在于人基因组序列中距离调控的靶基因 1～30 kbp 的启动子转录活性 DNA 序列，该序列可通过特有的染色质高级结构及其特异性蛋白因子的结合作用接近被调节的靶基因，促进基因转录，这种天然的距离调节模式是维系细胞内特异性发育阶段基因表达的结构基础。在人群中发现的一类 δβ-地贫，以成人期 Hb F 异常持续升高为主要特征，其分子缺陷的基础为基因的大片段缺失，包括成人期表达的人体主要成分-β-珠蛋白基因的丢失。有趣的是，β-基因缺失却使本已关闭的 γ-基因重新开放，并替代 β-基因的功能，其产物（γ-珠蛋白肽链）与 α-基因的产物（α-珠蛋白肽链）生成 HbF（$\alpha_2\gamma_2$）。图 10-6 为 3 种导致中国人 δβ-地贫的基因大片段缺失范围，在其 3′ 端缺失断裂点下游远端已经鉴定出某些特异性增强子序列，该增强子序列因大段丢失而被带到了邻近$^{G}\gamma$-基因的位点，此增强子可通过“距离效应”激活$^{G}\gamma$-基因，其结果是使$^{G}\gamma$-基因开放，并上调 γ-链的表达水平，实现红细胞特异性的 γ-基因替代 β-基因缺失的生物学效应。这一“增强子距离效应”能很好地解释上述两种不同缺失型（云南型和中国型）δβ-地贫之间表型上的差异。虽然这两种杂合子均有相似水平的 Hb F，但它们的纯合子之间却存在着显著不同。中国型有严重贫血表现（依赖输血维持生命），而云南型仅有轻度贫血，很显然，后者有着更强的促$^{G}\gamma$-基因表达的效应，究其分子基础，中国型缺失断裂点的 3′ 端位于云南型 3′ 端断裂点的下游约 10 kbp，这段缺失区中就包含了上面提到的增强子，中国型 δβ-地贫纯合子因不存在增强子效应，故表现为重度贫血。通过基因重排的位置效应使基因激活也是肿瘤发生的重要机制之一。

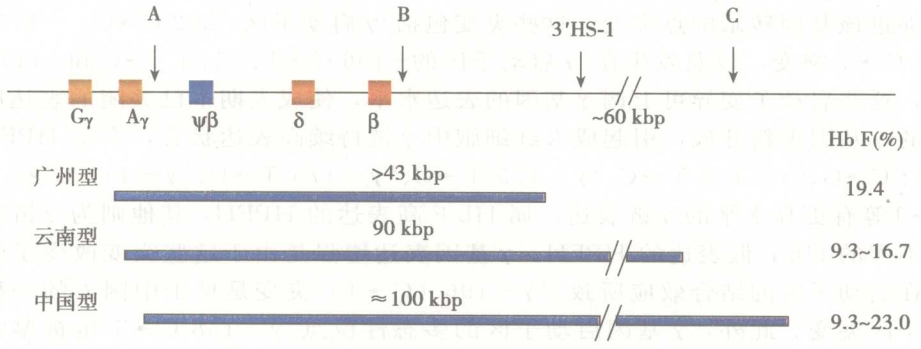


图 10-6　三种导致中国人 δβ-地贫的基因大片段缺失

图上部为人 β 类珠蛋白基因簇的组织结构示意图，其中功能基因用红色方框表示，假基因用蓝色方框表示，位于 β 基因 3′端下游的红系特异性的 LCR 中的高敏位点（3′HS-1）用向下箭头指出；A/B/C 向下箭头所指为 3 个增强子位点。图下部为有关病理基因的缺失范围，其左侧为缺失类型名称，右侧为各型的表型特征

（三）剂量效应

基因拷贝数变异是使基因功能增强的主要机制之一。神经系统常见的腓骨肌萎缩病（Charcot-Marie-Tooth disease，CMT）中的 CMT1A 型就是一种由基因重复（duplication）而导致的人类遗传性疾病。该病为常染色体显性遗传，其临床特征为足内侧肌和腓骨肌进行性无力和萎缩，其基因位于染色体 17p11.2-12，该基因编码一个 22 kDa 的周围神经髓鞘蛋白 22（peripheral myelin protein 22，PMP22），主要分布在髓鞘雪旺细胞膜，其功能可能与维持髓鞘结构的完整性、调节细胞的增殖有关。大多数 CMT1A 是由于 *PMP22* 基

因发生重复突变所致。正常人细胞中一个单倍体只有两个 *PMP22* 基因拷贝，而 CMT1A 患者则从亲代遗传了含 3 个甚至 4 个 *PMP22* 基因拷贝数的突变基因，该病也可由于新发突变所致，大部分新发突变患者为父源性 *PMP22* 基因的同源重组。基因重复可导致 *PMP22* 基因过度表达（基因剂量效应）可通过破坏细胞中高尔基氏体的功能，阻碍这一蛋白质的成熟所必需的加工步骤，因此而改变了髓鞘的组成，并使施万（Schwann）细胞参与髓鞘的发育及修复的功能失调，故引起髓鞘脱失和髓鞘再生。少数 CMT1A 还可由基因点突变所致，引起 CMT1A 的 *PMP22* 基因点突变并非通过上调基因表达水平而致病，而是影响蛋白质结构的错义突变。错义突变产生的突变蛋白会抑制其加工步骤，未加工好的蛋白质可损伤施万细胞的正常功能，在高尔基氏体形成的蛋白质被错误加工就可能导致髓鞘的不稳定而脱髓鞘。一般来说，*PMP22* 基因突变使得氨基酸发生改变会导致比基因重复更严重的症状。

基因的剂量效应也是解释肿瘤发生机制的基本理论基础之一。在肿瘤发生过程中，肿瘤抑制基因由于突变而失活，常可通过“去抑制”效应而使某些基因表达上调，细胞某些原癌基因的拷贝数异常增加（基因扩增），或由于基因突变（基因重复、倒位、插入和缺失等）激活受累位点的某些基因，也可导致细胞内相关基因剂量增加而使细胞持续分裂而发生癌变。如在小细胞肺癌细胞株中就有 *L-MYC*、*N-MYC* 和 *C-MYC* 基因拷贝数的扩增，其中尤以 *C-MYC* 的扩增更为明显，其拷贝数增加了数十倍甚至 200 倍之多。这类由基因扩增导致的基因剂量效应是肿瘤细胞的生物学活性特征之一。

（四）SNP 创造新启动子

目前认为，人基因组中广泛分布的 SNP，包括发生在基因编码区的中性替代在内，绝大多数点突变并不直接影响基因编码功能，但一些有潜在调节功能的 SNP 正在被揭示。例如，研究人员在调查 α-地贫的病因时，就发现了一种由调节 SNP（regulatory SNP，rSNP）导致人类遗传病的新机制，这一病理学机制的核心是 SNP 变异创造了一个新的转录启动子。被研究的对象是一组 α-地贫特征患病个体，经对 α-珠蛋白基因簇进行广泛分子筛查后，排除了导致 α-地贫表型的经典缺失或编码区点突变，且该基因簇的甲基化模式也无异常发现。进一步仔细观察后发现，病例样品 DNA 序列与正常对照组的重要区别是在 α-珠蛋白基因与其远端上游调节元件之间有一个 SNP 变异位点（位于 α-珠蛋白基因上游 15 kbp）。该 SNP 变异只是被测 DNA 区段 17 个 SNP 中的一个，病例样品此 SNP 位点为 G，而正常对照为 A。研究人员鉴定出此位点为一个功能获得性的 SNP 突变，即该位点的突变增加了一个与转录因子 GATA-1 结合的位点（TAATAA→TGATAA，后者为 GATA-1 结合位点），从而创造出一个新的转录启动子。新的启动子与原有的内源 α-珠蛋白启动子发生竞争，干扰了原有的内源性启动子活性，使其下游的 α-珠蛋白基因的表达显著下调，导致 α-地贫发生（图 10-7）。这一病理学机制的阐明为人类遗传病的鉴定提供了新的思路。需要说明的是，这一 SNP 突变的直接生物学效应属于功能获得性突变，既增加了相关位点 DNA 序列的转录因子结合功能，但如果从 α-地贫这一遗传病的编码基因的功能看，其效应是使 α-珠蛋白基因失去了功能。

（五）获得性 RNA 堆积

RNA 在细胞内异常堆积的致病机制可以很好地解释两种由不同致病基因缺陷引起的具有共同临床表型的强直性肌营养不良症的分子基础。强直性肌营养不良症（myotonic dystrophy，DM）为多系统受累的常染色体显性遗传性疾病，其临床特征为骨骼肌强直并发多器官疾病（如白内障、心律失常、心血管畸形、胰岛素抗性糖尿病、性腺发育不良等）。DM1 致病基因是位于 19 q13.2 的 *DMPK* 基因，该基因编码含 582 个氨基酸残基的萎缩性肌强直蛋白激酶。另一种近肢端肌营养不良症（proximal myotonic myopathy，PROMM）被归类为

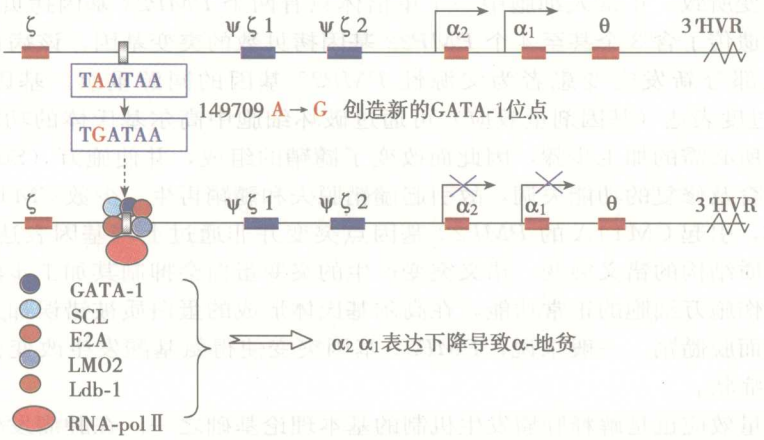


图 10-7　SNP 创造新启动子的分子机制

图上部为人 α 类珠蛋白基因簇的组织结构示意图，其中功能基因用红色方框表示，假基因用蓝色方框表示，位于 α 基因 3′端下游的高变区（HVR）用折线指出。图下部为各种参与 GATA 位点结合的转录调节因子。α 珠蛋白基因抑制作用╳表示

DM2，其致病基因是位于 3q13. 3-q24 的锌指蛋白 9 基因（*ZNF9*）。*DMPK* 基因全长 14 Kb，含 15 个外显子和 14 个内含子。致病突变为位于该基因 3′-端非翻译区（UTR）的一个 CTG 三联体重复扩增，正常人的（CTG）n 的拷贝数在 5～37 之间，而强直性肌营养不良患者增加为 50～3 000。*ZNF9* 基因的致病突变为该基因内含子 1 中的四核苷酸（CCTG）n 重复扩增，正常人（CCTG）n 重复次数最多为 26 次，而 DM2 患者（CCTG）n 重复范围极大，从 75 次到 11 000 次以上，平均 5 000 次。DM1 和 DM2 的核苷酸异常扩增均位于基因的 UTR，可以转录成 RNA，但不能翻译成蛋白质，杂合子个体除检测到扩增的异常转录本外，有功能的正常转录本亦同时存在，单倍型不足可以部分解释临床表型的发生机制。

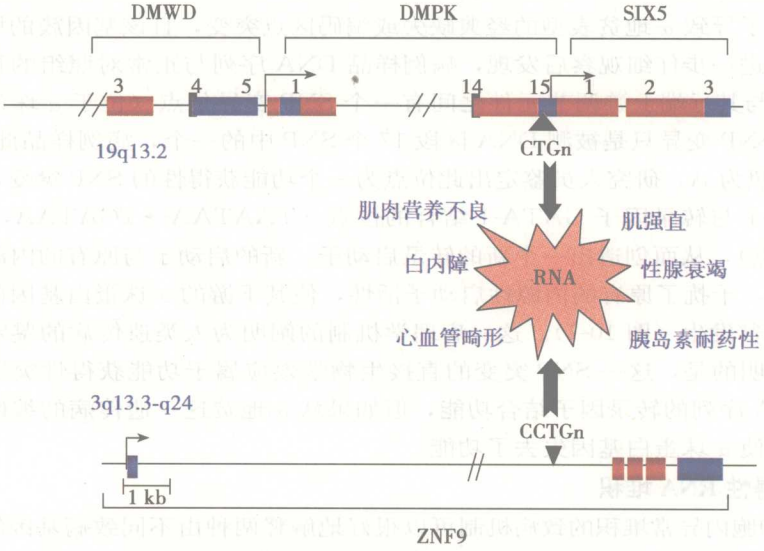


图 10-8　DM1 和 DM2 致病机理的 RNA 堆积共同通路

图上下方为人 DMPK 及其邻近基因和 ZNF9 基因的组织结构示意图，其中功能性基因外显子用红色方框表示，非翻译调节区和非编码基因外显子用蓝色方框表示，CTG 三联替扩增位于 DMPK 基因 3′UTR 区，CCTG 扩增位于 ZNF9 基因内含子 1 中。两个不同染色体位点的突变通过 RNA 堆积改变选择剪接效应等而引起肌肉和其他器官病变

研究证明 RNA 堆积在 DM 的发病机制上起着重要的作用，骨骼肌细胞或其他受累细胞核内含 CUG 和 CCUG 扩增子的异常转录本（RNA）的大量堆积是 DM1 和 DM2 致病机制的共同通路（图 10-8）。RNA 大量堆积的细胞毒性作用及其致病性有以下几方面：①堆积的 RNA 抑制成肌细胞（myoblast）的分化，使肌细胞发育障碍；②堆积的 RNA 直接损伤肌细胞导致进行性肌萎缩；③CUG-结合蛋白介导反式显性效应（trans-dominant effects）。当含 CUG 和 CCUG 扩增子的 RNA 堆积时，可使一组与之结合的可变剪接调节蛋白（如 CUG-结合蛋白）在细胞核内的含量增加，并改变其在细胞内的定位，其结果是，CUG-结合蛋白等反式作用因子参与的可变剪接调节功能失常，其下游靶点前体 RNA 的正常剪接被破坏，因而产生功能异常的剪接体。这一异常剪接机制是直接导致 DM1 和 DM2 病理表型的重要原因，如骨骼肌细胞膜上的氯通道蛋白的异常剪接体直接产生肌强直；心肌细胞肌钙蛋白 T 蛋白的异常剪接体与心律失常和心脏畸形有关；胰岛素受体的异常剪接体可诱导 DM 患者的胰岛素敏感性降低（胰岛素耐药）等。这一机制很好地解释了该病骨骼肌强直并发多器官疾病的临床表型。与 DM 类似的 STR 扩增（动态突变）引起的人类遗传病还有不少，这里介绍的 RNA 介导的病理机制对理解同类疾病有很好的启示。

上述两类产生相反生物学效应（基因功能的丧失或获得）的基因变异，及其各自不同的致病机制的阐述，在解释人类遗传性疾病发生的分子病理学的同时，也为我们理解人类基因的工作及其调控的基本原理提供了范例。

从基因表达的分子生物学基本概念上讲，无论是何种致病机制或细胞内异常代谢通路，基因突变对于基因功能的影响，其本质性生物学效应可以理解为以下两方面：①突变使受累基因所表达的蛋白质发生剂量改变（剂量不足或过量），从而影响其功能。影响基因转录水平（转录抑制或激活）的事件最为常见。以启动子为例的顺式元件的突变是这类变异的典型代表，其他有增强子突变和转录因子基因变异等。部分结构基因的剪接突变、无义突变和移码突变因导致基因表达产物迅速降解可降低基因剂量。部分三联体重复扩增使基因表达抑制。大片段突变中 DNA 缺失使基因丢失而无功能，以及基因重复使基因过量表达和导致基因转录激活的基因重排等也属此类基因变异；②蛋白质结构的改变，此类变异可伴有或不伴有基因表达水平的改变，即突变使受累基因所表达的蛋白质发生结构改变（一级结构导致高级结构改变），从而影响其功能。以错义突变为例的结构基因突变为这类变异的典型代表，其他还有部分结构基因的剪接突变、无义突变、移码突变和三联体重复扩增因可产生结构异常的基因表达产物，也可归为此类基因变异。

经典的人类遗传病的基因突变及其致病机制模式，主要围绕结构基因变异影响基因转录功能来展开说明，这是目前我们积累的与人类遗传病发生相关的主要知识，也是本章的主体内容所在。随着基因组学的研究进展，基因表达调控的知识不断被丰富和深化，基因变异在 mRNA 水平（如转录加工、稳定性及降解、核转运和翻译调控等）以及在细胞中的基因网络调控水平上对基因表达产物剂量及其结构影响的相关致病机制，目前虽然尚有待更系统的研究和积累，但已经研究阐明的一些新知识，是我们需要认真学习和掌握的。我们在本章给出的一些主题，如“反义 RNA 转录/基因组修饰”、“转录因子基因变异”、“SNP 创造新启动子”和“获得性 RNA 堆积”等是基于对这一领域的一些重要知识和概念的理解，人类遗传病致病机制及其基因工作原理的独特性和多样性由此可见一般。对基因变异的生物学功能的更深入的理解，还有赖于未来的研究中在细胞和分子水平上对基因及其调节机制的深入、全面的了解，而以人类疾病为对象的细胞和分子病理学机制研究，将会为我们提供越来越丰富的人类基因的分子生物学基础知识。

（徐湘民　熊　符）

第十一章　细胞异常增生性疾病的分子机制

关键知识点

- 编码产物可以促进肿瘤发生与发展的基因称为癌基因；反之称为抑癌基因。癌基因的活化或抑癌基因的失活是肿瘤发生的重要原因。单一癌基因或抑癌基因的突变所致肿瘤发生的几率较小，数种突变的积累最终将导致肿瘤发生、发展与转移。
- 细胞癌基因活化的机制包括被插入了额外的启动子或增强子、基因突变、基因扩增、基因重排及染色体易位等。癌基因编码产物一般是细胞增殖的正调控信号。
- 抑癌基因的编码产物是一些具有诱导细胞分化、维持基因组稳定、参与 DNA 损伤修复、控制细胞周期进程、触发衰老或诱导细胞程序性死亡等作用的蛋白质。重要的抑癌基因 *RB* 和 *P53* 的编码蛋白都是控制细胞周期转折点的重要分子。抑癌基因可因突变、缺失、丧失表达能力等改变而失活。

关键词： 癌基因；病毒癌基因；细胞癌基因；抑癌基因

正常细胞的增殖受到细胞内外多种信号分子的调控。一类属正调控信号分子，促进细胞的生长与增殖，并阻碍细胞的终末分化；另一类属负调控信号分子，抑制细胞增殖，促进细胞分化、衰老或凋亡。这些信号分子的调控失衡将引起异常的增殖和持续的分裂。

第一节　肿瘤细胞增殖机制

肿瘤是人类主要致死疾病，其发生是一复杂的生物学过程。目前认为，肿瘤的发生发展是一些编码细胞增殖、分化、凋亡调控信号以及 DNA 损伤修复蛋白的基因发生改变，导致细胞增殖调控失衡的结果。这些基因的改变有的来自于遗传，有的是基因与环境相互作用所致。其编码产物能够促进肿瘤发生与发展的基因被称为癌基因（oncogene）；其编码产物能够抑制肿瘤发生与发展的基因被称为抑癌基因（tumor suppressor gene）。

肿瘤细胞增生的核心问题是基因突变。如前所述，未能保真修复的 DNA 损伤效应之一是细胞发生基因突变。在基因水平上促使细胞增生及凋亡失去平衡，导致细胞恶性变（图 11-1）的原因有：①控制细胞增殖的原癌基因活化或抑癌基因失活；②促细胞凋亡的基因失活或抑制凋亡的基因功能增强；③DNA 修复基因失活，使突变在细胞内积累，并且累及到调节细胞增生及细胞凋亡的基因。这些基因中某一基因的单一突变所致肿瘤发生的几率较小，数种相关基

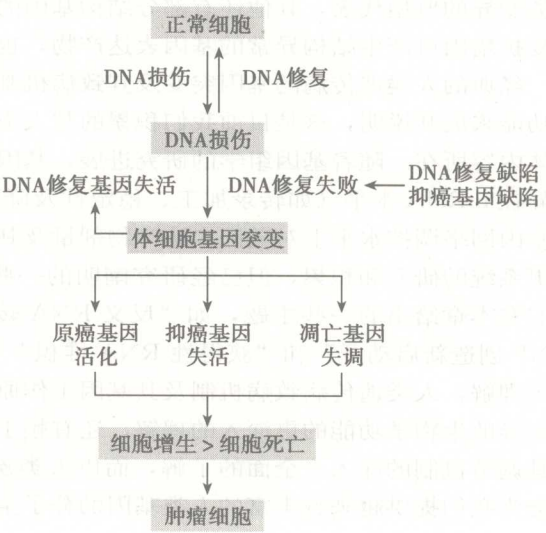


图 11-1　促进正常细胞向肿瘤细胞转化的因素

因突变的积累最终将导致肿瘤发生、发展与转移。

一、癌基因

癌基因是指细胞内或病毒内存在的、编码产物能促使正常细胞恶性转化（transformation），即发生恶性变的基因。存在于反转录病毒中的称为病毒癌基因（virus oncogene，v-onc）；存在于细胞中的称细胞癌基因（cellular oncogene，c-onc）或原癌基因（proto-oncogene）。

（一）病毒癌基因

癌基因最早发现于反转录病毒中。1911 年，美国洛克菲勒研究所的 F. Rous 医生将鸡肉瘤组织匀浆后的无细胞滤液皮下注射于健康鸡，发现可以引起肉瘤。20 世纪 50 年代确定了引起该种肉瘤的是一种病毒，命名为罗氏肉瘤病毒（Rous Sarcoma Virus，RSV）。从鸡肉瘤中分离得到的 RSV 在体外能使鸡胚成纤维细胞转化，在体内能使鸡患肉瘤。1975 年，M. Bishop 等比较了具备转化和不具备转化特性的 RSV 的基因组，发现野生型 RSV 中存在着一个与病毒生活史无关，但是能够转化鸡胚成纤维细胞，并使鸡患肉瘤的基因 *src*。以后又在其他反转录病毒中陆续发现了一些具有在体外转化细胞，在体内使宿主患肿瘤的基因，将这一类基因称为病毒癌基因，目前已发现的反转录病毒中的癌基因有 30 多种。

（二）细胞癌基因

缺失了癌基因的反转录病毒仍然能够正常地完成其生命周期，说明癌基因并不是它固有的必需基因。病毒中的癌基因从何而来？考虑到反转录病毒的生活史中的病毒基因组整合在宿主细胞基因组中的环节，人们推测病毒中的癌基因可能起源于宿主细胞。用核酸分子杂交的方法果然在正常的宿主细胞中找到了与病毒中的癌基因的同源基因，称为原癌基因或细胞癌基因。它们是病毒癌基因的源头。如猴肉瘤病毒中癌基因用 *v-sis*（simian sarcoma）表示，相应的细胞癌基因是 *c-sis* 表示。

（三）病毒癌基因与细胞癌基因的关系

目前认为病毒癌基因可能来源于宿主细胞癌基因，是前病毒整合在宿主细胞基因组后获得的（图 11-2）。细胞癌基因是断裂基因，有内含子，而相应的病毒癌基因均无内含子。推测其原因是由于整合在宿主细胞基因组中的前病毒获得细胞癌基因后在基因表达的 RNA 水平经过 RNA 转录后加工将内含子切除，因此包装在病毒颗粒中的病毒 RNA 基因组的病毒癌基因不再含有内含子。可以说，病毒癌基因原本不是病毒的基因，而是动物细胞正常基因组中的一个复本。当病毒在宿主细胞内复制时，由于 DNA 重组而将宿主细胞基因组中相当于病毒癌基因的序列重组到病毒的基因组内。所以说细胞癌基因是病毒癌基因的原型，又称为原癌基因。病毒癌基因对病毒本身无关紧要，却可使宿主细胞转化，引起肿瘤。

除此之外，病毒癌基因与相应的细胞癌基因比较，还存在着编码序列的点突变或缺失突变，因此表达的蛋白质功能有差别。病毒癌基因表达的蛋白质往往有较强的细胞转化活性。

（四）癌基因的功能

迄今发现的细胞癌基因都是一些具有重要功能的“看家基因”，在进化过程中高度保守。细胞癌基因编码产物的生理功能是调节细胞的生长和分化。生理条件下，细胞癌基因表达产物的活性和含量受到内外环境各种生长信号的严格控制，并不会导致细胞异常增

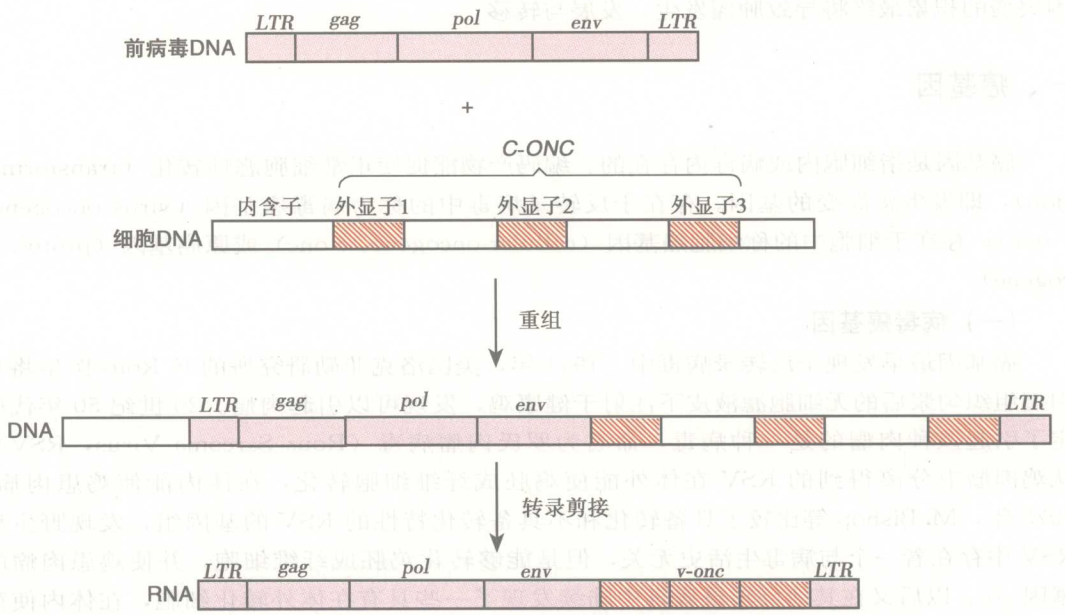


图 11-2　病毒癌基因来源于宿主细胞

殖。只有在突变后导致活性或含量过高时才会导致肿瘤发生，即发生癌基因的激活。

细胞癌基因按照结构特点有 *SRC*、*RAS*、*SIS*、*MYC*、*MYB*、*FOS*、*ERBA* 等基因家族，其表达产物属于各种生长因子、蛋白激酶、低分子量 G 蛋白、生长因子受体、转录因子等。表 11-1 列举了一些重要的细胞癌基因。

表 11-1　细胞癌基因及其产物举例

细胞癌基因	产物及其功能
c-*SIS*	血小板衍生生长因子（PDGF）β 链。PDGF 调节靶细胞正常生长与增生。相应的 v-*sis* 发现于猴肉瘤病毒中，产物为 $p28^{sis}$ 蛋白，能使靶细胞过度增生
ERB	编码 EGF 受体，与 EGF 结合后转导细胞增殖信号。相应的病毒癌基因 v-erbB 的产物为 $gp65^{erbB}$，缺失配体结合域，不需与 EGF 结合就可转导细胞增生信号
RAS 基因家族 *H-RAS* *K-RAS* *N-RAS*	产物的 Ras 蛋白属于小 G 蛋白，与 GTP 结合后有活性，能转导多种生长因子的细胞增殖信号。突变的 Ras 蛋白不具有 GTPase 活性，结合 GTP 后处于持续活化状态，导致细胞持续增殖
c-*SRC*	产物为 60 kDa 的胞浆内蛋白酪氨酸激酶，转导细胞增殖信号，v-*src* 产物能使细胞转化
c-*MYC*	产物为 49 kDa 的核内转录因子，与 Max 蛋白形成异二聚体，与特异的顺式元件结合，活化靶基因的转录。其靶基因多编码细胞增殖信号分子
Cyclin D 基因及 CDK4 基因	产物分别为细胞周期蛋白 D 和 CDK4，两者结合成复合物，调节细胞从 G_1 期过渡到 S 期。它们的活化可导致细胞转化

（五）细胞癌基因活化的分子机制

细胞癌基因在物理、化学及生物因素的作用下发生突变，表达产物的质和量的变化，表达方式在时间及空间上的改变，都有可能使细胞转化。从正常的原癌基因转变为具有使细胞转化功能的癌基因的过程称为原癌基因的活化，是功能获得的过程（gain of function）。细胞癌基因活化的主要机制如表 11-2 所示。

表 11-2　细胞癌基因活化机制

活化机制	举　例
基因自身突变	*c-RAS* 的点突变使之丧失 GTP 酶活性，持续活化
基因扩增	原癌基因通过基因扩增增加基因拷贝数，导致产物过量表达，使细胞发生转化。在肿瘤中表达量可因基因扩增升高几十甚至上千倍不等，例如小细胞肺癌中 *c-MYC* 的扩增
DNA 重排/染色体易位	白血病的异常基因 *BCR-ABL* 是 *ABL* 基因发生染色体易位所产生，导致 Abl 的蛋白酪氨酸激酶活性持续增高 *TRK* 与非肌原肌球蛋白基因间的重排导致产生持续活化的 TRK 激酶 伯基特（Burkitt）淋巴瘤细胞的染色体易位 t（8；14），致使 *c-MYC* 基因插入到免疫球蛋白重链基因的强启动子和增强子区下游，因而表达异常增加
病毒基因启动子及增强子的插入	禽类白细胞增生病毒（ALV）整合在禽类基因组中，由前病毒的长末端重复（LTR）序列中的启动子及增强子来调控宿主细胞内 *c-MYC* 表达，导致肿瘤产生

不同的癌基因有不同的激活方式，一种癌基因也可有几种激活方式。例如 *c-MYC* 的激活就有基因扩增和基因重排两种方式，很少见到 *c-MYC* 的突变；而 *RAS* 的激活方式则主要是突变。

上述各种细胞癌基因活化机制并未表明哪一种活化机制是使细胞转化的足够条件。事实上在肿瘤细胞中常发现两种或多种细胞癌基因的活化，例如白血病细胞株 HL-60 中有 *c-MYC* 和 N-*RAS* 的同时活化。实验也证明癌基因的协同作用可使细胞转化。例如原代培养的大鼠胚胎成纤维细胞传代 50 次左右就会死亡，导入重排的 *c-MYC* 可使它永生化，但细胞表型无恶性行为；如仅导入活化的突变 *RAS* 基因，细胞形态改变，但不能无限传代及在实验动物中形成肿瘤。不过如果同时导入 *c-MYC* 和 N-*RAS*，细胞就会发生恶性变，并在动物中成瘤。两种或更多的细胞癌基因活化可有协同作用，细胞癌基因的活化与抑癌基因的失活也产生协同作用。

二、抑癌基因

抑癌基因是一类可抑制细胞生长并有潜在抑癌作用的基因，当其受阻抑、失活、丢失、或其表达产物丧失功能可导致细胞恶性转化；反之，若导入或激活它则可抑制细胞的恶性表型。

对于一个特定组织或细胞恶性肿瘤的抑癌基因，需满足以下三个条件：①在癌的相应正常组织中必须有正常的表达；②恶性肿瘤细胞中相应基因应有所改变，包括点突变，片段缺失或表达缺陷；③在该基因缺陷的恶性肿瘤细胞中导入野生型（正常）基因将部分或全部抑制恶性表型。

（一）抑癌基因的发现

抑癌基因的发现较癌基因晚。20 世纪 60 年代开始的杂合细胞致癌性研究中将肿瘤与正常细胞融合，或在肿瘤细胞中导入正常细胞的染色体，都可获得无致癌性的杂合细胞，提示正常的细胞中有抑制肿瘤的基因，即抑癌基因。

A. Knudson 在研究视网膜母细胞瘤（retinoblastoma，RB）的流行病学中发现婴幼儿所患的肿瘤常常是双侧多发性，并有家族史；而没有家族史的幼儿往往发病年龄较大，而且是单侧性肿瘤。他假定遗传性和散发性的肿瘤都只与一个基因（即后来命名为 *RB* 的基因）有关，而且肿瘤的发生需要该基因位点发生两次突变。在早发性双侧肿瘤，患者从父

母获得的一对等位 *RB* 基因，其中有一个是正常的（野生型）基因，另一个是失活的。由于是胚系突变，患儿的体细胞中都有同样的 *RB* 基因型，因此只要野生型 *RB* 基因发生突变失活，*RB* 基因功能就会丧失，导致视网膜母细胞瘤产生。在迟发的单侧瘤，患者从父母获得的一对等位 *RB* 基因都有功能，只有它们都各自都突变失活才可能使 *RB* 基因的功能丧失而生长肿瘤。

用 DNA 限制性内切酶片段长度多态性（restriction fragment length polymorphism，RFLP）分析技术（见第二十章）研究发现，一些视网膜母细胞瘤细胞的染色体有 13q14 缺失。患儿正常视网膜母细胞中 13 号染色体长臂的 RFLP 分析表明它存在来自双亲的一对等位基因，而肿瘤细胞不再出现相应的 RFLP，称为杂合性丧失（loss of heterozygosity，LOH）。在视网膜母细胞瘤中，LOH 丢失的总是 *RB* 基因的野生型等位基因，使遗传获得的隐性突变性状得以显现。RFLP 研究表明，双侧视网膜母细胞瘤患者的正常视网膜母细胞中的一对 *RB* 基因中，一个是野生型的，另一个是突变型的；而单侧肿瘤患者的正常视网膜母细胞中有一对野生型 *RB* 基因，视网膜母细胞瘤中的一对 *RB* 基因都是突变的。如果将野生型 *RB* 基因经基因转染导入视网膜母细胞瘤细胞，转染后的细胞再接种动物后，不在动物中形成肿瘤，表明 *RB* 基因是抑癌基因。

（二）抑癌基因的功能

抑癌基因编码产物的功能主要有诱导细胞分化、维持基因组稳定、触发诱导细胞程序性死亡等。总体上抑癌基因对生长起着负调控作用，能抑制细胞的恶性生长。若抑癌基因发生突变，则不能表达正常产物，或者其编码蛋白产物的活性受到抑制，使细胞增殖调控失衡，导致肿瘤发生。

癌基因活化后导致肿瘤发生，而抑癌基因则相反，是在失活时导致肿瘤发生。抑癌基因失活的原因有基因突变、基因缺失、基因表达受抑等。抑癌基因产物的功能多种多样，目前已鉴定的几种抑癌基因产物及其功能如表 11-3。

表 11-3　某些抑癌基因及其功能丧失后导致的相关肿瘤

抑癌基因	染色体定位	产物细胞定位	功　能	相关肿瘤
DPCD4	18q21.1	细胞表面	转导 TGF-β 信号	胰腺癌、结肠癌等
PTEN	10q23.3	胞膜内侧	磷脂类信使的去磷酸化，抑制 PKB 的激活	恶性胶质瘤、膀胱癌、前列腺癌、子宫内膜癌等
NF1	17q11.2	胞膜内侧	抑制 Ras 信号转导	雪旺氏细胞瘤
NF2	22q12.2	细胞骨架	抑制 Rac 信号转导	脑膜瘤等
WT1	11p13	细胞核	转录因子	肾母细胞瘤（Wilms 病）、横纹肌瘤、肺癌、膀胱癌、乳癌、肝母细胞瘤
APC（*FAP*）	5q21-22	细胞质	抑制信号转导	胃癌、结肠癌等
RB	13q14	细胞核	负调节细胞周期	视网膜母细胞瘤、肺癌、骨肉瘤等
P53	17q12-13.3	细胞核	细胞周期负调节和 DNA 损伤后凋亡	大多数癌症
P16	9p21	细胞核	负调节细胞周期	胰腺癌、食道癌等
P21	6p21	细胞核	抑制 Cdk1、2、4 和 6	前列腺癌
NM23	17q22	细胞质	负调控 G 蛋白信号，影响微管聚合以调节细胞运动	乳腺癌、胃癌、结肠癌、直肠癌
DCC	11p15.5	细胞膜	细胞黏附	大肠癌、胰腺癌等
MEN-1	11q13	未定	TGF-β 信号有关	多发内分泌肿瘤

NF：neurofibromatosis 神经纤维瘤；APC：adenomatous polyposis of the colon 多发性结肠腺癌

（三）抑癌基因作用机制举例

1. *RB* 基因　*RB* 基因位于 13q14，全长 200 kbp，含 27 个外显子。*RB* 基因失活不仅与视网膜母细胞瘤及骨肉瘤有关，在许多散发性肿瘤，如 50%～85%的小细胞性肺癌、10%～30%乳腺癌、膀胱癌和前列腺癌中都发现了 *RB* 基因失活。

RB 基因在各种组织中普遍表达，产物是位于细胞核内的 105 kDa 的蛋白质 Rb^{105}。pRb^{105} 只有一条肽链，肽链中部有一个可折叠成口袋状的 AB 口袋结构域（AB pocket domain），它能与一些病毒的蛋白质及细胞蛋白质结合，肽链中还有可磷酸化修饰的位点。Rb 的磷酸化状态及与其他蛋白质的结合与它的功能密切相关。

细胞进入 G1 期时 Rb 处于低磷酸化状态，而低磷酸化的 Rb 使得细胞不能通过 G1-S 关卡。只有在细胞增殖信号通过依赖于周期蛋白 D1 的激酶 CDK4 的活化导致 Rb 蛋白磷酸化后，高磷酸化的 Rb 方允许细胞从 G1 期进入 S 期。因此低磷酸化的 Rb 在 G1 期特异的磷酸化是细胞从 G1 期进入 S 期的关键。低磷酸化的 Rb 对细胞周期的负调节作用是通过与转录因子 E2F-1 的结合而实现的（图 11-3）。低磷酸化的 Rb 的口袋结构域能结合 E2F-1 使之失活，S 期必需的基因产物如二氢叶酸还原酶、胸苷激酶、DNA 聚合酶 α 等合成因而受限，细胞周期的进展受到抑制。而高磷酸化的 pRb 不能与 E2F-1 结合，将导致这些基因的开放，促进细胞通过 G1-S 关卡。*RB* 基因的缺失使得细胞丧失了该关卡的“守卫”，细胞周期进程失控，细胞异常增生。

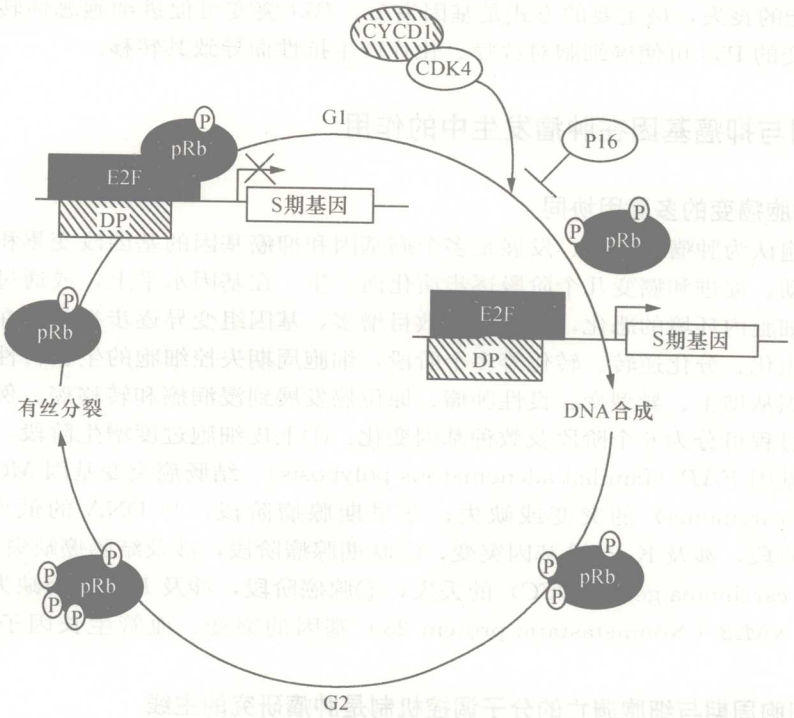


图 11-3　pRb 的磷酸化与细胞周期

2. *P53* 基因　是目前研究最多的、也是迄今发现在人类肿瘤中发生突变最广泛的抑癌基因。50%～60%的人类各系统肿瘤中发现有 *P53* 基因突变。某些家族遗传性肿瘤中还发现有胚系 *P53* 突变。无论是胚系 *P53* 突变或体细胞 *P53* 突变，常常是一对等位基因中只有一个等位基因有错义突变，造成 P53 蛋白中单个氨基酸残基替换。突变的 P53 蛋白不仅自身失去功能，它还能与野生型等位基因表达的 P53 蛋白聚合成无功能的四聚体，即显性

负突变（见第十章）。在某些肉瘤及一些淋巴瘤中，*P53* 基因的突变常常是等位基因双缺失、基因重排或剪接错误，导致 P53 蛋白缺失。

P53 基因位于 17p13.1，全长 20 kbp，含有 11 个外显子，编码一 53 kDa 肽链，活性形式为同源四聚体。P53 蛋白是位于细胞核内的一种转录因子，在各种组织中普遍存在。野生型 P53 蛋白半衰期很短，细胞内含量低。细胞受到射线辐射或化学试剂作用导致 DNA 损伤时，P53 蛋白水平升高，其原因主要是 P53 蛋白半衰期的延长及 P53 蛋白活化所致。例如，P53 蛋白的丝氨酸残基磷酸化，可使 P53 蛋白稳定性增加。

活化的 P53 蛋白的 N 端可以与转录辅激活因子 P300/CBP 结合，促进靶基因转录，其中一个重要的靶基因编码 P21 蛋白。P21 蛋白是 G1 期特异的细胞周期抑制物（见第九章），其作用是阻止细胞通过 G1-S 关卡点，使其停留于 G1 期。另一靶基因是生长阻滞及 DNA 损伤修复诱导基因 *45*（growth arrest-and DNA damage-inducible gene，*GADD45*），其产物是 DNA 损伤修复蛋白。P21 蛋白与 Gadd45 蛋白的共同作用使 DNA 受损的细胞不再分裂，并且修复损伤而维持基因组的稳定性。

P53 的另一个功能是促进细胞凋亡，当 DNA 损伤发生在已通过 G1-S 关卡点的细胞时，P53 蛋白可促进细胞促凋亡基因的表达，如 *BAX* 基因及 *FAS* 等，增加细胞凋亡。

P53 蛋白的最终效应是阻止 DNA 受损细胞通过 G1 期，而使细胞有足够的时间修复损伤的 DNA；如果损伤不能修复，则诱导细胞凋亡，从而避免产生具有癌变倾向的基因突变的细胞。起着稳定基因组和抑制突变细胞产生的作用，从而达到抑制肿瘤发生的目的。P53 正常功能的丧失，最主要的方式是基因突变，*P53* 突变可促进细胞恶性转化，减少细胞凋亡。突变的 P53 可使癌细胞对放疗、化疗产生抗性而导致其转移。

三、癌基因与抑癌基因在肿瘤发生中的作用

（一）细胞癌变的多基因协同

目前普遍认为肿瘤的发生、发展是多个癌基因和抑癌基因的基因改变累积的结果，经过起始、启动、促进和癌变几个阶段逐步演化而产生。在基因水平上，或通过外界致癌因素，或由于细胞内环境的恶化，突变基因数目增多，基因组变异逐步扩大；在细胞水平上则要经过永生化、分化逆转、转化等多个阶段，细胞周期失控细胞的生长特性逐步得到强化。结果组织从增生、异型变、良性肿瘤、原位癌发展到浸润癌和转移癌。例如，结肠癌的发生发展过程可分为 6 个阶段及数种基因变化：①上皮细胞过度增生阶段，涉及家族性多发性腺癌基因 *FAP*（familial adenomatous polyposis）、结肠癌突变基因 *MCC*（mutated in colorectal carcinoma）的突变或缺失；②早期腺瘤阶段，与 DNA 的低甲基化有关；③中期腺瘤阶段，涉及 K-*RAS* 基因突变；④晚期腺瘤阶段，涉及结肠癌缺失基因（delete in colorectal carcinoma gene，*DCC*）的丢失；⑤腺癌阶段，涉及 *P53* 基因缺失；⑥转移癌阶段，涉及 *NM23*（Nonmetastatic protein 23）基因的突变、血管生长因子基因表达增高等。

（二）细胞周期与细胞凋亡的分子调控机制是肿瘤研究的主线

1. 癌基因、抑癌基因与细胞周期　细胞周期调控体现在细胞周期驱动和细胞周期监控两个方面，前者在第九章已有详细讨论。细胞周期监测机制由 DNA 损伤感应机制、细胞生长停滞机制、DNA 修复机制和细胞命运决定机制等构成。细胞一旦发生 DNA 损伤或复制错误，将会启动 DNA 损伤应激机制（第十七章），经由各种信号转导途径令细胞停止生长，修复损伤的 DNA（第四章）。如果 DNA 损伤得到完全修复，细胞周期可进入下一个时相，正常完成一个细胞分裂周期；倘若 DNA 损伤修复失败，细胞凋亡机制将被启动，

损伤细胞进入凋亡，从而避免DNA损伤带到子代细胞，维持了组织细胞基因组的稳定性，避免肿瘤发生的潜在可能。

肿瘤细胞的最基本特征是细胞的失控性增生，而失控性增生的根本原因就是细胞周期调控机制的破坏，包括驱动机制和监控机制的破坏。监控机制破坏可发生在损伤感应、生长停滞、DNA修复和凋亡机制的任何一个环节上，结果将导致细胞基因组不稳定，突变基因数量增加，这些突变的基因往往就是癌基因和抑癌基因。同时，很大一部分的癌基因和抑癌基因又是细胞周期调控机制的组成部分。因此，在肿瘤发展过程中，监控机制的异常会使细胞周期调控机制进一步恶化，并导致细胞周期驱动机制的破坏，细胞周期的驱动能力异常强化，细胞进入失控性生长状态，从而细胞出现癌变性生长。

2. 癌基因、抑癌基因与细胞凋亡　细胞除了生长、增生和分化等现象之外，还存在细胞死亡现象，即程序性细胞死亡或凋亡。有些抑癌基因的过量表达可诱导细胞发生凋亡，而与细胞生存相关的癌基因的激活则可抑制凋亡，细胞凋亡异常和肿瘤的发生发展的密切关系才引起研究者的重视。现在已经明确细胞凋亡在肿瘤发生、胚胎发育、免疫反应、肿瘤免疫逃逸、神经系统发育、组织细胞代谢等过程中起重要作用。

总之，肿瘤分子生物学的进展已经深刻地改变了人们对肿瘤发生和生命现象的认识，并使肿瘤研究从以揭示肿瘤病因和寻找肿瘤治疗方法为目的单项研究，转变为以研究整个生命现象为目的，全面揭示细胞生命分子机制的重要手段。肿瘤分子生物学必将在整个生命医学研究中发挥越来越重要的作用。

第二节　器官纤维化的分子机制

机体器官的基本结构包括实质（parenchyma）和间充质（stroma）两部分，实质是指执行该器官功能的细胞，如肝细胞，肾曲管细胞，心肌纤维等。间充质是构架及支持功能的结缔组织、血管及神经纤维。结缔组织由大量细胞外基质，包括胶原、弹性纤维、网状纤维等纤维、纤连蛋白、层连蛋白等糖蛋白及各种蛋白多糖和较少的细胞，如成纤维细胞、巨噬细胞、浆细胞及脂肪细胞等组成。机体的组织由于各种原因受损后，若损伤不严重，该组织的实质细胞具有增生能力，可由实质细胞增生修复而完全恢复原状，若组织损伤比较严重，无法由实质细胞增生修复时，结缔组织替代实质细胞修复，最终导致纤维化（fibrosis）。因此纤维化是组织损伤后由结缔组织修复的结果。

1. 肾纤维化　各种原因引起的慢性肾疾病进展为终末期肾衰的实质是肾纤维化和硬化，从而导致有效肾单位丧失和肾功能的进行性下降。近年研究揭示，炎症细胞浸润，肾固有细胞在致纤维化性生长因子、细胞因子等作用下发生表型改变等是肾纤维化形成的重要机制。肾小管间质纤维化（tubulointerstitial fibrosis，TIF）是肾硬化的一个重要特征。无论原发病变在肾小球、肾小管还是在血管，大多数进展性肾病均可见到TIF，而TIF的出现又往往预示着病变预后不良。如果没有同时伴有小管、间质炎症性浸润及纤维化，肾小球硬化乃至肾功能不全就很少发生。肾小管细胞的作用十分重要，当其受损时，可以表达细胞黏附分子，释放趋化因子、细胞因子及生长因子、单核细胞趋化蛋白-1（monocyte chemotactic protein 1，MCP-1）、粒-单核细胞集落刺激因子等，这些因子可以将炎症细胞、单核细胞及淋巴细胞等吸引到肾间质。而单核细胞通过进一步地释放致纤维化性细胞因子，刺激成纤维细胞增生、胶原合成，从而促进TIF的形成。淋巴细胞产生的IL-2、γ-干扰素等可通过刺激巨噬细胞产生TGF-β而进一步促进间质的成纤维细胞产生胶原。肾小管细胞本身亦可通过释放IL-1、PDGF、TGF-β等而直接促进肾间质成纤维细胞的激活与增生。

2. 肝纤维化　肝纤维化是细胞外间质合成与降解的稳态平衡被慢性肝实质持续的炎症与坏死所破坏，细胞外间质合成大于降解，导致肝内纤维结缔组织异常增生的病理过程。在各种致病因素作用下，肝细胞变性坏死，导致趋化性/炎症性因子的释放，这些细胞因子可激活肝内的星状细胞并使之增殖。与此同时，被激活的库普弗细胞、窦内皮细胞、巨噬细胞等在坏死区释放 TGF-β 和 PDGF 等细胞因子，促使星状细胞转变成成纤维细胞。成纤维细胞能产生胶原、大分子糖蛋白和蛋白多糖，是产生基质的主要细胞，对肝纤维化的形成起主要作用。成纤维细胞作为静止期星形细胞的完全转化形式，通过自分泌和旁分泌机制，进一步促进静止期星形细胞增生和活化，另一方面分泌胶原、非胶原蛋白、蛋白多糖等细胞外基质。细胞外基质大量生成，在坏死的肝细胞周围有新的成纤维细胞包围，基质合成增加导致瘢痕形成。活动性基质产生及正常或新生基质的降解减少，使得基质在肝内沉积，从而导致肝纤维化。在此过程中，基质的生成又会反过来影响细胞的发生、分化、增殖、黏附和表型的表达。

3. 肺纤维化　特发性肺纤维化（idiopathic pulmonary fibrosis，IPF）病因不明，发病机制亦未完全阐明，但已有足够证据表明与免疫炎症损伤有关。综合近年来的研究，关于 IPF 发病机制及其过程可概括如下：①未知抗原激活 B 细胞，产生抗体并形成免疫复合物，进而刺激和活化肺泡巨噬细胞；②活化的肺泡巨噬细胞释放各种炎性介质直接损伤肺细胞及基底膜结构，另外增加了与纤维化形成密切相关的分子，如纤连蛋白（fibronectin）、肺泡巨噬细胞原性生长因子（alveolar macrophage-derived growth factor）、PDGF 和 IGF 等的合成与释放，它们能吸附成纤维细胞，并刺激其增殖，并介导胶原基质收缩；③在肺泡巨噬细胞释放的 IL-8、TNF-α 等介导下中性粒细胞向肺泡聚集，形成以中性粒细胞比率增高（20%）为特征的肺泡炎，而中性粒细胞炎症反应又释放一系列介质，引起或加重肺损伤与纤维化；④成纤维细胞增生和产生胶原是该病的重要环节和结局。正常人成纤维细胞生长存在精确的调控，如前列腺素 E2、成纤维细胞移动抑制因子等均属于负调节因子。另在 IPF 发现一种编码 PDGF 的 c-*SIS* 基因，与转移性病毒癌基因 v-*sis* 非常相似，因此 IPF 的发生是否代表了成纤维细胞的负调节失效，或成纤维细胞的“肿瘤性”增生，是十分有兴趣的问题。IPF 患者肺间质胶原的合成速度或总量并无增加，但Ⅰ型胶原增加、Ⅰ型胶原对Ⅱ型胶原的比率升高。由于Ⅰ型胶原是一种高张力强度、低顺应性、呈平行排列的交叉带状纤维，它的增加足以解释 IPF 的形态和生理学改变，而无论胶原总量增加与否。

（韩跃武）

第十二章　感染性疾病的分子生物学

关键知识点

- 宿主对细菌和病毒等病原体的感染和患病易感性取决于病原体基因表达产物与宿主的相互作用。
- 外毒素是病原菌的主要致病因子，进入细胞后可以作用于宿主细胞内一些重要的功能调控分子，直接或间接地影响宿主细胞的膜结构及信号转导、基因表达等过程，使宿主细胞的正常功能受损或丧失。细菌可能通过主动或被动地获得对抗菌物质的耐药性，并可在个体间散播。
- 不同病毒以特定的方式侵入宿主，在宿主的易感细胞中进行复制和基因表达的过程称为感染。宿主细胞基因组可编码表达特异的病毒受体，介导病毒的吸附结合、穿入、脱壳、病毒增殖、装配与释放等过程，完成对细胞的感染。病毒侵入机体后在宿主细胞中的复制会对宿主造成损伤，这种损伤可能是病毒复制并破坏细胞直接导致的结果，也可能是宿主的免疫系统识别受感染细胞后做出的过度反应引起，或者二者兼而有之。

关键词： 感染性疾病；致病岛；外毒素；抗药性基因；病毒受体

感染性疾病（infectious diseases）是指由病原微生物（如病毒、细菌、衣原体、支原体、立克次体、螺旋体、放线菌、真菌）和寄生虫等病原体侵入机体所导致的疾病。感染性疾病包括具有传染性的，即过去称为传染病（communicable diseases）和一些传染性不强的非法定传染病的感染性疾病，以及无传染性的自体正常寄居的微生物引起的感染。不同病原体感染所致疾病的发病机制各不相同，这里不一一阐述，本章主要讨论病原菌致病基因和病毒致病基因及其作用机制。

第一节　病原菌致病的分子机制

病原微生物基因组中的致病基因，亦称毒力（virulence）基因，是病原菌致病的结构和功能基础。了解这些基因及其表达产物的结构和功能特点，对病原微生物的哪些基因与毒力有关、哪些与其在宿主体内定居有关、哪些与持续感染有关等进行详细研究将有助于理解其致病分子机制，从而为相关疾病的防治提供依据。

一、病原菌的基因组

自 1995 年美国基因组研究所完成流感嗜血杆菌的全基因组序列起，目前已完成了超过 100 种（株）病原菌和 2 000 多株病毒的全基因序列测定，目前已完成的部分常见病原微生物基因组情况见表 12-1。

病原微生物基因组的研究意义在于：①加速发现新的重要致病基因。以流感嗜血杆菌为例，在获得其完整基因组前，只有 7 种与内毒素合成相关的基因被发现。而在获得其基因组后，人们很快就发现和鉴定了 25 个新的致病相关基因。②寻找特异的病原 DNA 序列，提高临床诊断的效率和准确性。菌株特异性 DNA 序列可用于分型，而特异的毒力基因则可用于预测和判断治疗效果。③筛选有效药物和发展疫苗，找到细菌生存所必需的基因以及细菌感染过程中的关键基因。

表 12-1　常见致病细菌的基因组

病原菌株	所致疾病	基因组大小（Mb）	病原菌株	所致疾病	基因组大小（Mb）
生殖道支原体 G237	尿道炎	0.58	奈氏脑膜炎球菌 B 型 MC58	脑膜炎	2.27
肺支原体 UAB CTIP	呼吸道疾病	0.96	金黄色葡萄球菌 N31525Hu50	皮肤感染、败血症	2.81
沙眼衣原体 D/ UW-3/CK	沙眼	1.05			
沙眼衣原体 Mopn	沙眼	1.07	李斯忒单细胞菌 EGD-2-e	李斯忒病、流产	2.95
普氏立克次 Madrid E	斑疹伤寒	1.1	白喉棒状杆菌 NCTC13129	白喉	3.1
苍白螺旋体 Nicholas	梅毒	1.14	分枝麻风杆菌	麻风	3.1
肺炎衣原体 CWL029	呼吸道疾病	1.23	霍乱弧菌 N16961	霍乱	4.03
伯氏疏螺旋体 B31	莱姆病	1.44	分枝结核杆菌 H37Rv	结核病	4.4
胎儿弯曲菌 NCTC11169	胃肠炎	1.64	大肠埃希菌 K-12 MG1655	肠炎	4.6
幽门螺旋杆菌 J99	胃溃疡	1.64	耶森鼠疫杆菌 CO92	鼠疫	4.65
幽门螺旋杆菌 26695	胃溃疡	1.66	福氏志贺菌 2a301	痢疾	4.7
嗜血流感杆菌 Rd	脑膜炎、中耳炎	1.83	伤寒沙门菌 CT18	伤寒	4.8
奈氏脑膜炎球菌 A 型 Z2491	脑膜炎	2.18	铜绿假单胞杆菌（绿脓杆菌）PA01	条件致病	6.3
肺炎链球菌 TIGR4	肺炎	2.2			

对病原细菌基因组研究的结果表明，染色体中许多编码细菌毒力的基因通常集中在一段独特的 DNA 片段上，称为毒力岛或致病岛（pathogenicity islands，PAI），PAI 两侧通常具有重复序列和插入序列，这说明毒力岛通常分布在可移动的遗传元件上。病原菌基因组中一般还含有特定的耐药基因，细菌的耐药基因有的是固有耐药基因；有的则是获得性耐药基因，后者要么是由基因突变获得，要么是细菌间通过质粒转化获得。

二、病原菌毒素及其编码基因

细菌能引起机体产生疾病的性能称为致病性或病原性（pathogenicity）。致病菌的致病作用与它的毒力、侵入数量以及侵入途径密切相关，而其中毒力主要指侵袭力和毒素，是细菌致病的关键。大多数毒力因子由质粒或噬菌体携带，也可能存在于细菌的致病岛或转座元件中。

细菌在增殖过程中，可以合成并分泌导致宿主细胞损伤的毒性物质，称为外毒素（exotoxin）。外毒素多为热不稳定的蛋白分子，往往通过结合于细胞表面的特殊分子后进入宿主细胞，并对宿主细胞的某些关键功能分子进行修饰，使之丧失功能并最终杀灭细胞。相对于外毒素，内毒素（endotoxin）则是致病菌（一般是革兰阴性菌）的菌体破裂后释放出来的毒性物质，其成分为脂多糖（lipopolysaccharide，LPS）。内毒素的致病作用将在第十三章讨论，这里仅讨论外毒素的致病机制。外毒素基因可由染色体 DNA、质粒或噬菌体编码。常见细菌外毒素基因及编码产物特性见表 12-2。

表 12-2　常见细菌外毒素基因及编码产物特性

基因位置	细菌名称	毒素名称	基因及编码产物的特征
染色体 DNA	霍乱弧菌 *Vibrio cholerae*	霍乱毒素 *cholera toxin*，CT	由 1 个 A 亚基和 5 个 B 亚基构成，A、B 亚基分别由 *ctx* 操纵子中 A、B 结构基因编码。A 亚基又有 A_1 和 A_2 两个组分，其中 A_1 是毒素的毒性组分，借助 A_2 和 B 亚基结合；B 亚基与小肠黏膜上皮细胞膜上 GM_1 受体结合，导致毒素分子变构，A 亚基与 B 亚基解离进入细胞
	痢疾志贺菌 *Shigella dysenteriae*	志贺毒素 *shiga toxin*，ST	由一个 A 亚基和 5 个 B 亚基组成，A、B 亚基分别由染色体上的 *stx* A 和 *stx* B 编码。其中 B 亚基与宿主细胞膜上 Gb_3 结合，帮助毒性的 A 亚基进入细胞
	幽门螺杆菌 *Helicobacter pylori*，HP	空泡毒素 *vacuolating cytotoxin A*，VacA	由 *vacA* 基因编码，存在信号序列区（s 区）和中间区（m 区）。s 区和 m 区以不同形式编码的 5 种嵌合体，即 s_{1a}/m_1、s_{1a}/m_2、s_{1b}/m_1、s_{1b}/m_2 和 s_2/m_2。其中 s_{1a}/m_1 型菌株毒力最大，而 s_2/m_2 型菌株的毒力最小
		细胞毒素基因相关蛋白 A *cytotoxin associating gene protein A*，*CagA*	由致病岛上的 *cagA* 基因编码。*cag* 致病岛是长约 40 kbp 的插入序列，分为 *cag* Ⅰ区和 *cag* Ⅱ区，*cagA* 结构基因位于Ⅰ区，上游存在细胞因子诱导基因 a 和 b
质粒 DNA	炭疽杆菌 *Bacillus authracis*	炭疽毒素 *anthrax toxin*，AT	由保护性抗原（protective antigen，PA）、水肿因子（edema factor，EF）和致死因子（lethal factor，LF）构成。编码三个组分的基因均位于炭疽杆菌中的 pOX1 质粒上
	肠产毒性大肠埃希菌（*Enterotoxigencic E. coli*，*ETEC*）	肠毒素 *enterotoxin*	不耐热肠毒素（heat labile enterotoxin，LT）属于 A-B 结构模式毒素，由 1 个 A 亚基和 5 个 B 亚基组成，其致病机制与霍乱毒素相似；耐热肠毒素（heat stable enterotoxin，ST）为低分子量多肽
噬菌体	白喉棒状杆菌 *Corynebacterium diphtheriae*	白喉毒素 *diphtheria toxin*，DT	细菌本身无产毒基因，DT 是由 β-棒状噬菌体（*corynephage β*）携带的 *tox* 基因编码，当 β-棒状杆菌噬菌体侵袭白喉杆菌后，在溶原阶段，*tox* 基因整合到宿主染色体，成为毒性白喉棒状杆菌。DT 属于 A-B 结构模式的毒素，由 1 个 A 亚基和 1 个 B 亚基组成
	肠出血性大肠杆菌 *Enterohemorrhagic Escherichia coli*	志贺样毒素 *shiga-like toxin*，SLTs	SLT 的结构基因 a 和 b 位于噬菌体上。SLT 与志贺毒素结构相似

三、细菌毒素以各种方式影响宿主细胞功能

外毒素进入宿主细胞后，可以作用于细胞内一些重要的调控分子，直接或间接影响宿主细胞的膜结构、信号转导或基因表达等过程，使宿主细胞功能受损或丧失。表 12-3 列举了部分常见细菌外毒素的作用方式和致病机制。

表 12-3　常见细菌外毒素的作用方式和致病机制

作用方式	毒素名称	毒素来源	作用部位	致病机制
改变宿主细胞通透	大肠杆菌溶血素 A *E. coli hemolysin A*	多种 G^- 菌	肠道黏膜	N-端具有保守的 10 个疏水 α-螺旋，可在红细胞膜形成疏水性孔道，使红细胞受到胶体渗透压休克而溶解
	肉毒毒素 *botulinus toxin*，BT	肉毒梭菌	神经组织	结合于神经末梢突触前膜，通过胞饮作用进入胞质，阻塞突触前膜阳离子通道，抑制突触后膜的乙酰胆碱释放
干扰信号转导	炭疽毒素 AT	炭疽杆菌	吞噬细胞	其中的水肿因子是腺苷酸环化酶，可增加胞内 cAMP 浓度，促进 IL-6 基因的表达。致死因子具有蛋白水解酶活性，可水解 MAPKK 或 MEK 的 N-端使之灭活。此外，LT 可诱导巨噬细胞大量表达 IL-1 和 TNF-α
	霍乱毒素 CT	霍乱弧菌	肠道黏膜	可作用于 G 蛋白的 αs 亚基，使之发生 ADP-核糖化修饰，抑制其自身 GTPase 活性，从而持续激活 Gsα，最终造成 Cl^- 通道开放，细胞在释放 Cl^- 的同时伴随大量的失水
	肠毒素	大肠杆菌	肠道黏膜	其中热稳定性毒素 α 的裂解产物可激活内源性鸟苷酸环化酶，促进 cGMP 的生成，最终引起 Cl^- 的释放
	百日咳毒素 *pertussis toxin*，PT	百日咳鲍特菌	呼吸道	PT 作用与霍乱毒素相似，对 G 蛋白进行 ADP-核糖化修饰，但其作用靶点是 Gi，因而最终引起细胞内 cAMP 水平下降
影响基因表达	志贺毒素 ST	志贺菌	小肠黏膜	A 亚基为 RNA 糖苷酶，可以使核糖体 28S RNA 失活，从而抑制蛋白质合成
	白喉毒素 DT	白喉棒状杆菌	多器官	B 亚基可以特异地对肽链延长因子 2（EF-2）进行 ADP-核糖化修饰，从而抑制宿主细胞的蛋白合成

细菌的致病因子除外毒素外，黏附性结构物质，如菌毛、黏附性蛋白及其他成分也在致病过程中起重要作用。

四、抗药性基因的出现使细菌产生耐药

天然或人工合成的抗菌物质（包括抗生素及抗菌肽等）可以迫使细菌发生遗传进化改变，产生耐药性。细菌对抗菌物质产生耐药的机制主要有以下几种：

（一）产生抗生素的钝化酶或灭活酶

某些细菌在突变、接合或转化过程可能获得特异的基因，它们编码某些水解酶或钝化酶，能够特异地对某些抗生素进行水解或修饰，使抗菌药物在作用于菌体前即被破坏或失效。

钝化酶或灭活酶基因通常位于质粒 DNA，耐药质粒广泛存在于革兰氏阳性和革兰氏阴性细菌中，可通过多种途径在细菌之间传播耐药基因。例如有些耐药质粒编码 β-内酰胺酶，也称灭活酶，β-内酰胺酶可催化青霉素类和头孢菌素类抗生素分子结构中的 β-内酰胺

水解，失去抗菌活性。耐药质粒编码的氯霉素乙酰转移酶和氨基乙酰化酶可以分别对氯霉素和氨基糖苷类抗生素进行乙酰化修饰从而使之钝化失活。转座子也是常见的传递耐药基因的遗传物质，转座子的两端侧翼序列是两个反向重复序列，含有转座酶基因和耐药基因等。携带耐药基因的转座子可在细菌的染色体、质粒和噬菌体基因组之间转移，导致耐药基因的播散。

（二）靶蛋白基因突变

抗菌药物对细菌的作用均有相应的作用靶点，它们很多是蛋白质，如代谢关键酶、功能蛋白或核糖体蛋白等。如果这些靶蛋白基因出现变异，则抗菌药物不能再对这些靶点起作用或作用减弱，从而产生耐药性。如细菌的二氢叶酸合成酶突变可产生对磺胺的耐药。具有甲氧西林（meticillin）抗性的葡萄球菌可获得一种突变的青霉素结合蛋白 PBP2a，其与甲氧西林的亲和力比正常 PBP 低得多，从而可以对一定浓度的抗生素产生抵抗。具有链霉素抗性的细菌主要是核糖体相关蛋白基因突变而致，如位于 16S RNA 的 *rrs* 基因突变或是核糖体蛋白 S12 的 *rpsL* 基因突变。结核杆菌的耐药相关基因突变包括对异烟肼特异的 *katG* 突变、*inhA*/*mabA* 和 *ahpC* 突变，对利福平特异的 *rpoB* 突变，对氟喹诺酮特异的 *gyrA* 和 *gyrB* 突变等。

（三）改变细菌细胞壁的通透性

细菌通过改变细胞膜或细胞壁的通透性，抑制抗生素进入细胞或降低抗生素在细胞内的浓度来产生耐药性。例如当革兰阴性菌的非特异性孔道蛋白（porin）基因发生突变而使其表达量降低时，由这些孔道介导的抗菌药物的流入则减少，从而增加了对抗菌药物的耐受性。某些具有高渗透性外膜的细菌往往对抗菌药物敏感，它们可以通过形成生物菌膜来降低外膜通透性产生耐药。绿脓杆菌对多种常用抗生素耐药，除因为其外膜存在着独特的药物泵出系统外，还由于其外膜蛋白 Opr F 使药物不易通过。

（四）主动外排机制

某些细菌可以利用流出泵（efflux pump）的特殊主动转运机制将进入细菌内的抗菌药物转运出去，例如耐药的革兰氏阳性和革兰阴性菌均存在由 Ter 膜蛋白介导的四环素泵出系统。大肠杆菌的多药耐药机制是由其膜上的主动外排蛋白 AcroA、AcroB 和外膜通道蛋白 TolC 协同完成的。AcroB 是三个亚基形成的水母状的同源三聚体，具有 3 个结构域：①跨膜域由 3 个亚基的 12 个 α-螺旋构成，锚定在细菌的内膜；②孔道域（pore domain）

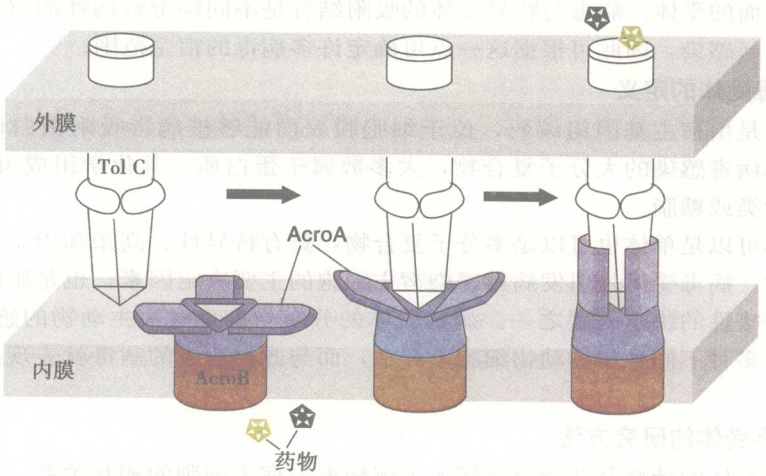


图 12-1　AcroA 和 AcroB-TolC 系统介导的药物主动排除模式

由 3 个亚基的 β-α-β 模体构成，孔道域构成中空结构，孔道的下端开口于孔道腔，上端连接 TolC 对接域，孔道腔内有通道与周浆间隙相通；③TolC 对接域（TolC docking domain）是 3 个 TolC 对接域形成的漏斗状结构，其顶端开口，TolC 蛋白对接于此开口。依靠 AcroAB-TolC 系统，细胞内的物质通过跨膜域或周浆间隙的物质经通道易位到孔道腔，并经 TolC 管主动转运出细菌外膜（图 12-1）。AcroAB-TolC 系统主动外排的物质范围较广，可为阳性、中性和阴性物质。例如，该系统直接参与了铜绿假单胞菌对 β-内酰胺类抗生素的耐药。

第二节　病毒致病的分子机制

在病原微生物引起的感染性疾病中，由病毒引起的约占四分之三。人类的许多常见感染性疾病，如肝炎、脑炎、肺炎、脊髓灰质炎、流行性感冒、狂犬病、艾滋病等，都是由病毒引起的。

病毒是一类具有独特生物学性状的非细胞型微生物，具有个体微小、结构简单、专性寄生且仅在活细胞内复制增殖、耐冷怕热，对干扰素敏感等特点。病毒基因组只含有一种核酸（DNA 或 RNA），病毒进入宿主细胞后，利用宿主细胞的各种酶和蛋白因子进行复制和外壳蛋白的合成，再装配成子代病毒颗粒，并进一步在宿主体内扩散，感染其他细胞。在病毒致病机制的研究中，主要集中在病毒基因产物对感染细胞的毒性作用、宿主对病毒基因表达产物的各种应答反应、病毒基因对宿主细胞基因的作用（如整合、抑制或激活部分基因）等方面。这里我们将从病毒进入细胞的机制和病毒感染对宿主细胞功能的影响两方面对其进行阐述。

一、病毒通过病毒受体进入宿主细胞

病毒在细胞外处于静止状态，只有进入活细胞后才能发挥其生物活性。这是由于病毒不具备细胞器及代谢必需的酶系统，因此决定了它的专性寄生特性。病毒必须侵入易感的宿主细胞，依赖宿主细胞的原料、能量、酶及蛋白因子来复制病毒的核酸和表达病毒蛋白，病毒的这种特殊增殖方式称为“复制”。

病毒增殖的第一步是病毒表面成分（包膜、衣壳蛋白或一部分纤突）特异地吸附结合于宿主细胞表面的受体。病毒与特异受体的吸附结合是不同病毒致病性的重要基础，不能结合就不会引起感染，因此可根据这一点可确定许多病毒的宿主范围。

（一）病毒受体的定义

病毒受体是由宿主基因组编码，位于细胞膜表面能够被病毒吸附蛋白识别并与之结合，从而引起病毒感染的大分子复合物，大多数属于蛋白质，其化学组成可能是糖蛋白、蛋白聚糖、脂类或糖脂。

病毒受体可以是单体也可以是多分子复合物，具有特异性、高亲和力、饱和性及独特的生物学活性。病毒受体是引发病毒感染宿主细胞的主要决定因素，也是影响病毒宿主特异性和组织亲嗜性的决定因素之一。病毒受体的分布与病毒对宿主动物的感染范围有关，有的受体可在多种不同种属的动物细胞上存在，而与此相适应的病毒就表现出较广的宿主范围。

（二）病毒受体的研究方法

研究病毒受体的特性及功能对于深入了解病毒与宿主细胞的相互关系，在分子水平上阐明病毒感染的致病机制具有重要意义，同时也为研制更有效的病毒疫苗、抗病毒药物和

诊断试剂，甚至对新病毒的分类等提供帮助。

归纳起来，目前用于确定病毒受体的方法包括：①利用人工合成的特异性肽段与病毒受体进行竞争性结合或封闭细胞表面的病毒受体；②利用病毒受体的单克隆抗体阻断病毒与受体的结合；③利用特定酶的作用对病毒受体上的特异成分进行降解；④亲和层析分离、纯化病毒受体或病毒-受体复合物；⑤利用抗独特型抗体模拟病毒来寻找细胞受体蛋白；⑥用转基因方法使受体缺陷性细胞获得特异性受体；⑦利用噬菌体表面表达呈现（phage display）技术筛选病毒受体；⑧用酵母双杂交方法确定病毒与受体结合特性及相互作用的精确位点。

（三）目前已获得的部分病毒受体

近年来，各种病毒在人细胞表面的受体研究方面已经取得了较大进展，表 12-4 列举了部分已确认的病毒受体及辅受体。

表 12-4　部分病毒受体的特性

病毒	受体分子	受感染细胞
乙肝病毒（HBV）	IgA 受体	肝细胞
艾滋病毒（HIV）	CD4、趋化因子受体	T 细胞、巨噬细胞
麻疹病毒	CD46（补体受体）	白细胞、上皮细胞
EB 病毒	CD21（Ⅱ型补体受体）	B 细胞、树突状细胞
鼻病毒	CD54（黏附分子 ICAM-1）	淋巴细胞、单核细胞、上皮细胞、内皮细胞
ECHO 病毒	CD55（补体调节分子 DAF）	白细胞
	CD49b（黏附分子 β1 整合素）	淋巴细胞、单核细胞、血小板
巨细胞病毒	CD13	单核细胞、粒细胞
冠状病毒	CD13	单核细胞、粒细胞
脊髓灰质炎病毒	Ig 超家族	脊髓运动神经细胞
狂犬病毒	乙酰胆碱受体	横纹肌细胞
流感病毒 A 和 B 副黏液病毒	含唾液酸的糖蛋白或糖脂	红细胞，上皮细胞
流感病毒 C	含 9-O-乙酰唾液酸的糖蛋白或糖脂	上皮细胞

二、病毒感染宿主的基本过程

从病毒进入细胞，经过复制与基因表达到子代病毒的释放，称为一个复制周期。复制周期包括吸附、穿入、脱壳、生物合成及装配释放等五个步骤。

1. 吸附与结合　病毒与宿主细胞可由于偶然碰撞和静电作用而形成可逆疏松的结合，当宿主细胞表面存在特异的受体时，则可通过结构的互补性与病毒表面的受体连结蛋白（receptor binding protein，RBP）发生特异的、稳定的结合。

2. 穿入　穿入是指病毒核酸或感染性核衣壳穿过细胞膜进入细胞质的过程，主要有三种方式：①融合（fusion）病毒囊膜与细胞膜融合，核衣壳进入细胞质。副粘病毒类常以融合方式进入，如麻疹病毒、腮腺炎病毒囊膜上的融合蛋白带有一段疏水氨基酸，可介导细胞膜与病毒囊膜的融合。②病毒吞饮（viropexis）当病毒与受体结合后，在细胞膜的特殊区域与病毒一起发生内陷，形成膜性囊泡进入细胞质。病毒在细胞内仍被胞膜覆盖。某些囊膜病毒，如流感病毒借助病毒的血凝素（hemagglutinin，HA）完成与质脂膜间的

融合，囊泡内的低 pH 环境使 HA 蛋白的三维结构发生变化，从而介导病毒囊膜与囊泡膜的融合，核衣壳进入细胞。③直接进入　某些无囊膜病毒，如脊髓灰质炎病毒与受体接触后，衣壳蛋白的构象改变并对蛋白水解酶敏感，病毒核酸可直接穿越细胞膜进入到细胞质中，而大部分衣壳蛋白仍留在胞膜外，这种进入的方式较为少见。

3. 脱壳　病毒核酸从衣壳释放的过程称为脱衣壳或脱壳（uncoating），包括去除包膜、蛋白质骨架和病毒核酸外壳。不同病毒脱壳的方式不同，有的十分简单，如脊髓灰质炎病毒，在吸附穿入细胞的过程中病毒 RNA 即可直接释放到胞质中；有包膜病毒的脱壳过程往往较复杂，病毒在与细胞膜融合过程中首先脱去包膜（或外衣壳），然后再进一步脱去内衣壳。如痘苗病毒的内衣壳进入细胞后，病毒体要先发生多聚酶活化，合成病毒脱壳所需要的酶，才能最终完成脱壳。进入细胞的病毒衣壳可被吞噬体中的溶酶体酶降解而去除。

4. 病毒增殖　DNA 病毒和 RNA 病毒复制的生化机制有明显区别，但复制的结果都是合成病毒核酸分子和蛋白衣壳，然后装配成新的感染性病毒。病毒完成一个复制周期大约需 6～8 小时。根据病毒核酸类型的差异，病毒的复制又分为双股 DNA 病毒的复制、正链 RNA 病毒的复制、负链 RNA 病毒的复制、反转录病毒的复制等（见第三章）。

5. 装配与释放　新合成的病毒核酸和病毒结构蛋白在感染细胞内组合成病毒颗粒的过程称为装配。新的病毒颗粒从细胞内转移到细胞外的过程为释放。大多数 DNA 病毒，在宿主细胞核内复制 DNA，在细胞质内合成蛋白质，再转入核内装配成熟。病毒装配成熟后释放的方式有：①宿主细胞裂解，病毒释放到周围环境中，如腺病毒、脊髓灰质炎病毒等无囊膜病毒；②以出芽的方式释放，常见于有囊膜病毒，如疱疹病毒在核膜上获得囊膜，流感病毒在细胞膜上获得囊膜而成熟，然后以出芽方式释放出成熟病毒，也可通过细胞间桥或细胞融合感染邻近的细胞。

三、病毒感染对宿主细胞功能的影响

病毒感染宿主细胞后会对宿主的正常功能造成严重影响并引起损伤，这种损伤可能是病毒复制并直接破坏宿主细胞的结果，也可能是宿主的免疫系统识别受感染细胞后做出的过度反应引起，或者二者兼而有之。临床上表现出的各种症状，如发热、组织损伤、疼痛、恶心等都可能是机体对病毒感染做出反应的表现。

（一）病毒对宿主细胞的直接作用

病毒对宿主细胞的直接作用包括杀细胞感染、稳定感染和整合感染。由于病毒致病机制较复杂，因此在病毒感染和致病不同阶段分别发挥不同作用。病毒感染对宿主细胞功能的直接影响包括：

1. 直接损伤　某些病毒抗原成分插入细胞膜后，引起细胞膜抗原改变，使细胞融合或发生自身免疫性细胞损伤。另外某些病毒的毒性衣壳蛋白，如腺病毒的刺突可对细胞产生直接破坏。

2. 影响细胞大分子的合成　这是抑制细胞生长的关键，如单纯疱疹病毒（HSV）等合成的蛋白质能抑制和降解细胞 DNA 及 mRNA；脊髓灰质炎病毒能抑制 mRNA 的 5′-帽结构的形成，或产生大量病毒 mRNA 与宿主细胞 mRNA 竞争核糖体，从而阻断宿主细胞蛋白质的合成。另外，很多杀细胞病毒编码的早期蛋白能将宿主细胞合成大分子所需的各种酶类和蛋白因子据为己用，抑制甚至关闭宿主细胞 RNA 和蛋白质的合成，继而影响其 DNA 的合成，使细胞正常代谢紊乱，最终导致细胞死亡。

有些病毒在整合感染期，病毒 DNA 整合到宿主基因组时可影响宿主基因的正常表达，

这种影响既可能是抑制，也可能是激活，导致宿主基因表达失控。反转录病毒在复制过程中，其基因组 RNA 经反转录成为 cDNA 后，整合于宿主细胞染色体 DNA，由于反转录病毒基因组两端的长末端重复序列（long terminal repeat，LTR）中含有强启动子和增强子，因此当 LTR 插入到宿主细胞癌基因的启动子区域或邻近处，往往可以启动并促进癌基因的表达，最终导致细胞增殖和肿瘤发生。

3. 破坏膜功能　一些病毒感染细胞后，可破坏细胞骨架，导致细胞膜和细胞器膜的通透性改变，使细胞内环境的离子、营养物质平衡以及跨膜运输遭到破坏。某些病毒感染细胞后，产生的病毒或子代病毒颗粒也可对细胞内的其他膜性结构产生破坏性损伤，如破坏溶酶体引起细胞自溶。另一些病毒如副粘病毒、疱疹病毒和反转录病毒，还能介导感染细胞与邻近未感染细胞的融合，形成多核巨细胞，造成细胞功能的障碍。

4. 诱导或抑制细胞凋亡　许多病毒在感染细胞后都可以诱导细胞凋亡，如疱疹病毒、微小病毒、黏液病毒以及各种反转录病毒（如艾滋病毒）。当这些病毒感染宿主细胞后，病毒本身或由病毒编码蛋白间接地作为诱导因子引发细胞凋亡或被细胞毒 T 淋巴细胞诱导凋亡。另一方面，某些病毒则可抑制细胞凋亡，如腺病毒早期基因 E1B 的表达产物可通过抑制 *P53* 基因的表达而抑制凋亡；EB 病毒编码的 LMP-1 可诱导细胞表达 bcl-2 来抑制凋亡。病毒对细胞凋亡的抑制可以维持自身的生存与增殖，导致持续病毒感染和宿主细胞的过度增殖，并引发肿瘤。

（二）病毒感染的免疫病理损伤作用

在大部分病毒感染性疾病中，机体的损害并不是由于病毒的直接致细胞病变效应引起的，而是病毒抗原刺激免疫应答对机体造成的间接损伤，这种损伤称为免疫病理损伤，是机体为清除病毒而付出的代价。诱发免疫病理反应的病毒特异性抗原可以是病毒体表面的包膜或核衣壳（无包膜病毒）抗原，也可以是在病毒体内部的基质、核蛋白等。

1. T 淋巴细胞介导的病理损伤　病毒抗原与宿主细胞的Ⅰ类主要组织相容性抗原（MHC-Ⅰ）一起提呈给 $CD8^+$ 的杀伤性 T 淋巴细胞（cytotoxic T lymphocyte，CTL），使之活化。CTL 杀伤靶细胞的机制可能是释放胰蛋白酶样的丝氨酸蛋白酶和细胞毒性淋巴因子，如穿孔素（perforin）等，导致靶细胞溶解。$CD4^+$ T 淋巴细胞也可以被特异的诱导而激活，产生细胞因子和趋化因子，由此引发炎症反应。

2. B 淋巴细胞介导的病理损伤　当病毒复制迅速，免疫系统无法及时清除或无法到达病毒的感染部位时，病毒就会与特异性抗体结合形成复合物。病毒-抗体复合物会沉积在毛细血管中引起病变，若激活补体系统，会引起Ⅲ型变态反应，如关节炎、肾小球肾炎等。

3. 诱发自身免疫　有的病毒抗原与宿主组织抗原含有共同的抗原决定簇，当宿主的免疫系统对这些共同的抗原决定簇产生免疫反应时，就会引起自身免疫性疾病。目前较有力的支持证据是疱疹病毒感染引起的基质角膜炎。另外转基因动物实验证明，病毒感染后产生的干扰素等细胞因子可以活化无反应性 T 细胞，打破免疫耐受，这些活化的 T 细胞可释放某些细胞因子，破坏局部环境的平衡状态，从而导致自身免疫。此外，病毒也可刺激多克隆 B 细胞，产生抗细胞抗体。

4. 病毒超抗原诱发的免疫反应　一些病毒蛋白是非常有效的 T 细胞刺激物，称为超抗原（superantigen）。它们能结合抗原提呈细胞上的Ⅱ类 MHC 分子，然后直接激活 T 细胞，避开了将抗原降解为多肽再由 MHC-Ⅱ递呈给 T 细胞这一环节，从而缩短了抗原提呈路径。超抗原的活化能产生大量 T 细胞，将破坏免疫系统的协同性，从而引发多种疾病。超抗原大多是病毒产物，如狂犬病毒的核蛋白、巨细胞病毒及 HIV 病毒编码的特异蛋白质等。

第三节　典型感染性疾病的分子机制

感染性疾病的种类繁多，致病机制各异。现以临床上常见的几种典型感染性疾病为例，了解病原微生物的致病分子机制。

（一）霍乱的致病机制

霍乱弧菌（*Vibrio cholerae*）是霍乱（俗称 2 号病）的病原体。霍乱弧菌有两种生物型，即古典型（*classic biotype v. cholerae*）和埃尔托型（*El Tor biotype v. cholerae*）。而霍乱的“O”抗原血清型目前至少发现有 155 种，其中具有致病性的仅有 O1 群和 O139 群。

霍乱弧菌的致病因子有多种，包括菌毛、鞭毛、血凝素、肠毒素等，其中肠毒素，即霍乱毒素（cholera toxin，CT）是最主要的致病因子。

1. 鞭毛与菌毛　霍乱弧菌进入小肠后，依靠活泼的鞭毛运动穿过肠黏膜表面的粘液层，通过菌毛附着于肠壁上皮细胞刷状缘的微绒毛上，并在其上繁殖。霍乱弧菌的菌毛在电镜下观察至少有 A、B、C 三种类型，菌毛的功能可能是克服邻近细胞间的静电斥力，使菌体与肠细胞在相隔一定距离情况下，由菌毛搭桥而黏附在一起，从而不被肠蠕动所清除。菌毛与肠毒素有协同表达作用，故又称毒素协同表达菌毛，简称 TCP。霍乱弧菌鞭毛的化学组成为蛋白质，鞭毛轴心相对分子质量为 4.7×10^3 kDa，鞘蛋白相对分子质量为 4.9×10^3 kDa。

2. 霍乱毒素　霍乱毒素 CT 是目前已知最强的致腹泻毒素，是一种不耐热的毒素，对蛋白酶敏感。CT 由 1 个 A 亚基和 5 个 B 亚基组成，是典型的 A-B 型毒素。A 亚基含有毒性组分，与 B 亚基结合存在。B 亚基与小肠黏膜上皮细胞神经节苷脂（GM1）受体结合，

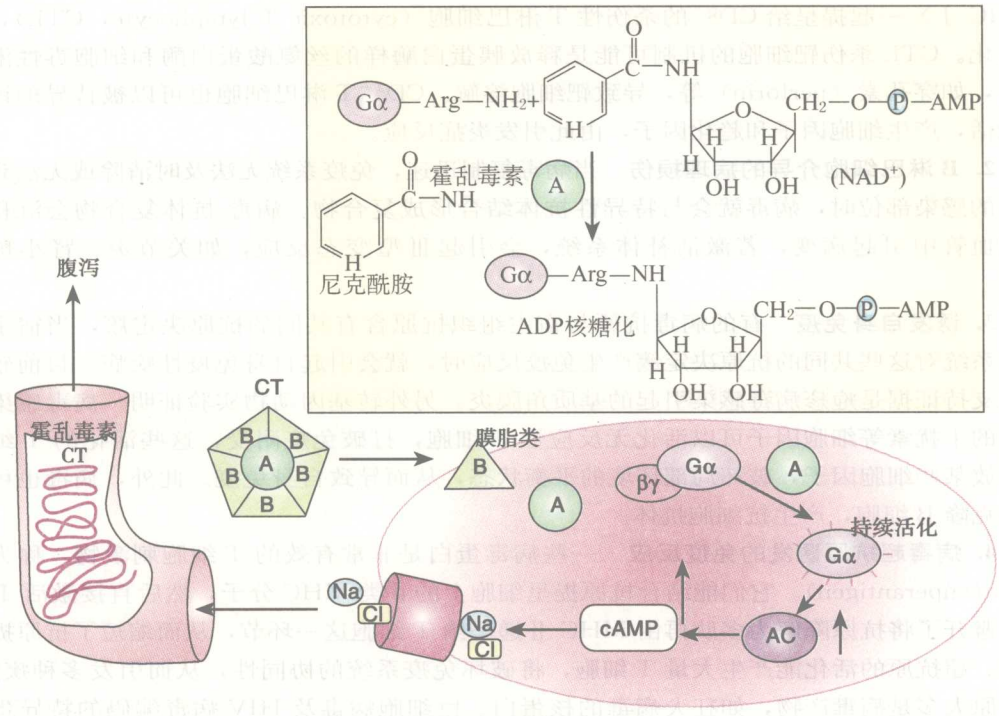


图 12-2　霍乱毒素的作用机制

导致毒素分子变构，从而使 A 亚基脱离 B 亚基进入细胞内。A 亚基具有 ADP-核糖基转移酶活性（图 12-2），可催化 NAD^+ 上的 ADP 核糖转移到激动型 G 蛋白（Gs）的 α 亚基上，使其失去内在的 GTPase 活性而被持续激活，后者激活腺苷酸环化酶（AC），使细胞内 cAMP 水平升高，最终使小肠黏膜上皮细胞膜上的 Cl^-/HCO_3^- 通道开放，细胞内离子外流的同时，伴随大量水的丢失，导致腹泻。

3. 其他致病因子　霍乱弧菌产生的神经氨酸酶可以增强霍乱毒素与上皮细胞结合的能力。O139 群除具有上述毒力因子和相关基因外，还有荚膜多糖和特殊的毒性决定簇，其功能是抵抗血清中杀菌物质和使细菌黏附到小肠黏膜细胞上。

（二）HIV 的致病机制

人类免疫缺陷病毒（human immunodeficiency virus，HIV）是一种发现相对较晚（1983 年）的病毒，属于反转录病毒科中的慢病毒属。HIV 感染引起免疫缺陷综合征（acquired immunodeficiency syndrome，AIDS），即艾滋病。AIDS 以传播迅速、致死性高及进行性免疫系统损伤为主要特征，是当今严重威胁人类健康的感染病之一，目前已知的致病机制主要有：

1. 损伤 $CD4^+$ T 淋巴细胞导致免疫应答障碍　CD4 分子是 HIV 的主要受体，当病毒与 $CD4^+$ T 淋巴细胞接触时，病毒包膜表面的糖蛋白 gp120 与细胞表面的 CD4 分子 V1 区结合，致使与其紧密相连的包膜跨膜蛋白 gp41 发生变构，其疏水端深入细胞膜内，促进病毒包膜与宿主细胞膜的融合和核衣壳的侵入。另外，除 CD4 外，HIV 对巨噬细胞和 T 细胞的感染还依赖某些趋化因子受体（如 CCR5、CCR2b、CCR3 及 CXCR4 等）作为辅受体的参与。已发现存在一些 HIV 健康携带者，其不发病的原因就是由于辅受体变异、病毒不能进入细胞。

HIV 感染 $CD4^+$ T 细胞后，能以较快的速度复制，导致此类细胞的病变和死亡，使 $CD4^+$ T 细胞数量减少和功能降低，$CD8^+$ T 细胞比例则相对增高，出现 $CD4^+/CD8^+$ 倒置，从而引起患者机体免疫功能的紊乱。

2. 其他细胞损伤　由于 CD4 分子也存在于其他细胞，如单核巨噬细胞、树突状细胞、小神经胶质细胞、郎格罕细胞、肾小管上皮细胞、纤维母细胞等，因此这些细胞亦对 HIV 敏感。

$CD4^+$ 细胞的 HIV 感染是 AIDS 患者发生免疫缺陷的前提。HIV 导致 $CD4^+$ 靶细胞损伤的机制有：①病毒的出芽释放和其包膜糖蛋白插入细胞膜，都能引起宿主细胞膜通透性的改变；②大量病毒增殖产生的未整合的 cDNA，干扰宿主细胞代谢；③HIV 的 gp120 与靶细胞的 MHC-Ⅱ有同源性，抗 gp120 抗体能与这类 T 细胞起交叉反应，引起同种异基因免疫应答，造成自身免疫损伤；④gp120 可介导多个 $CD4^+$ 细胞的融合，促进多核巨细胞的形成而导致细胞死亡；⑤机体对病毒感染的细胞产生特异性的 CTL 细胞毒性或 ADCC 效应；⑥病毒与易感细胞结合后，直接诱导细胞凋亡。

3. HIV 的变异　HIV 具有高度变异性，决定病毒致性的主要基因，包括 *env*、*nef*、*LTR* 及 *pol* 基因均有较高的变异率，但在不同病毒株间在上述基因的变异率各不相同。*env* 基因编码包膜糖蛋白，是变异率最高的基因，其变异与 HIV 的流行和逃避宿主的免疫应答密切相关。HIV 的高度变异性对疫苗的研制和 AIDS 的防治带来很大的困难。

（三）乙肝病毒的致病机制

乙型肝炎病毒（Hepatitis B virus，HBV）属嗜肝 DNA 病毒科，是一种部分环化的双链 DNA 病毒。HBV 可通过多聚人血清清蛋白受体、直接侵入和 IL26 介导等多种途径侵入肝细胞，导致肝细胞浊肿、毛玻璃样变性坏死等。

HBV 除可对肝细胞进行直接损伤外，更重要的是，HBV 与相应的抗体形成免疫复合

物介导机体产生的免疫应答反应可能才是其致病的主要原因。HBV 的致病机制主要包括：

1. 病毒介导的免疫病理损伤　HBV 感染时发生的肝细胞损伤是特异性和非特异性细胞免疫应答协同作用的结果，如特异的 CTL 介导的细胞毒活性、巨噬细胞释放 TNFα 引起的非特异性细胞毒作用以及 NK 细胞的非特异性细胞毒活性。

CTL 在急性 HBV 感染时对病毒清除起主要作用，但在 CTL 清除病毒的同时也造成肝细胞损伤。诱发细胞毒活性的抗原主要是 HBV 基因组的 c 基因所编码的核衣壳蛋白，即细胞型的 HBcAg 和分泌型的 HBeAg。HBcAg 的氨基酸序列与 HBeAg 有很大的同源性，都可以表达于感染细胞表面。感染肝细胞中的核壳寡肽与肝细胞膜表面的Ⅰ类人白细胞抗原（human leucocyte antigen Ⅰ，HLA-Ⅰ）分子结合，形成肽-HLA 复合物。CTL 表面的 CD8 分子和抗原受体共同识别肝细胞表面的 HLA-Ⅰ和核壳寡肽，诱发细胞自溶。

2. 病毒致机体免疫应答低下　HBV 感染后，抑制干扰素的合成，且使靶细胞的 HLA-Ⅰ类抗原表达降低，因此使依赖 CTL 的免疫应答反应减弱。此外，感染 HBV 后机体 IL-2 合成也会减少，这与 HBV 感染淋巴细胞有关。幼龄感染 HBV 后，因免疫系统尚未发育成熟，可对病毒形成免疫耐受，从而不出现或仅出现低的抗病毒免疫，病毒可长期存在于体内。

3. 免疫复合物引起的损伤　在 HBV 感染者体内常形成 HBsAg-抗 HBsAb 复合物，这些免疫复合物可沉积于肾小球基底膜和关节滑液等处，激活补体，引起Ⅲ型变态反应，其中以关节炎和肾炎最为常见。当大量免疫复合物沉积于肝细胞时，可以引起毛细血管栓塞，并诱导巨细胞释放肿瘤坏死因子 α（TNF-α），TNF-α 与靶细胞膜上的 TNF-α 受体结合后向靶细胞内移，被溶酶体摄取，使溶酶体的稳定性降低，多种水解酶外流引发细胞自溶，表现为急性肝坏死。

4. 诱发原发性肝癌　流行病学研究证实，HBV 与原发性肝癌存在密切关系。原发性肝癌的发生机制可能涉及多种因素，如病毒基因的整合、病毒的变异等。

（傅　强　覃　扬）

第十三章　炎症的分子机制

关键知识点

- 炎症是具有血管系统的生命体对损伤因子所发生的复杂防御反应，是损伤、抗损伤和修复三位一体的综合过程。炎症的诱发因素有生物因素、理化因素、组织坏死、超敏反应等。
- 生物致炎因素的受体主要是 TLR 家族的成员，理化致炎因素经由伤害性信号受体引起细胞反应。
- 导致炎细胞在炎症部位聚集的原因主要有两个：一是血流中的白细胞和血管内皮细胞表面的黏附分子间的相互作用；二是组织中趋化因子的吸引。
- 介导炎症反应的化学因子称为炎症介质，按来源可将其分为细胞释放的炎症介质和体液中的炎症介质。
- 炎症相关信号通路是一复杂的网络系统，其中 TLRs、TNF-α 受体、趋化因子受体等介导的信号转导通路是重要的炎症相关反应通路。
- 炎症反应可受到反馈调节。促使早期炎症反应终止的信号主要有脂氧素、腺苷、细胞间相互作用、胆碱能神经递质、TLRs 信号通路中的负性调节分子等。中枢神经系统也参与炎症的调节。

关键词： Toll 样受体；趋化因子；炎症介质；细胞因子；急性期反应蛋白；脂氧素

当外源和内源性损伤因素作用于机体后，细胞和组织可发生各种各样的损伤性变化，同时局部和全身也发生一系列复杂反应，以局限和消除损伤因素，清除和吸收坏死组织和细胞，并修复损伤，这种综合的机体防御反应称为炎症（inflammation）。炎症是具有血管系统的生命体对损伤因子所发生的复杂防御反应，是损伤、抗损伤和修复三位一体的综合过程。

炎症的诱发因素很多，概括起来主要有：①生物因素（细菌、病毒、寄生虫等）；②物理因素（低温、高温、创伤、射线等）；③化学因素，包括外源性化学物质（强酸、强碱、毒物等）和内源性化学物质（体内代谢产物等）；④缺血或缺氧等原因引起的组织坏死；⑤因机体免疫反应状态异常而引发的超敏反应。依据病理学特征，炎症反应可分为以渗出为特征的急性炎症反应和以增生为特征的慢性炎症反应。

本章将着重介绍炎症刺激信号及其受体、参与炎症反应的细胞及其作用机制、主要炎症介质及其作用机制、炎症反应相关的信号转导机制、炎症的急性期反应、炎症反应的调控点等内容，旨在概要阐明炎症发生、发展和转归的分子机制。

第一节　炎症刺激信号及其受体

细胞接受各种炎症刺激信号的分子多属膜受体，这些受体接收炎症刺激信号后，将信息转导入细胞内，细胞再释放出各种炎症介质，使组织炎症反应得以放大，以消除损伤性刺激因素。

一、生物性致炎因素的主要受体——TLR 家族

接受生物性致炎信号刺激的主要是存在于细胞的膜受体，其中最为重要的一类受体是

Toll 样受体（Toll-like receptor，TLR）家族。

至 2007 年 12 月，在人类已成功克隆 11 个 TLR 家族成员（TLR1-11），在小鼠已成功克隆 13 个 TLR 家族成员，TLR1～13，但 TLR10 不表达。TLRs 属Ⅰ型跨膜蛋白，膜外是富含亮氨酸的重复序列，胞内有一个负责信号转导的 TIR（Toll-IL-1 receptor）同源结构域。TLRs 的配体主要为不同的病原体成分，称为病原体相关的分子模式（pathogen-associated molecular patterns，PAMPs），这些 PAMPs 与哺乳类细胞的成分无同源性，但 TLR11 的配体例外。TLR11 的配体为鼠弓形体表达的一种 profilin 样蛋白（一类能引起超敏反应的过敏原），该蛋白在人和植物均有同源物。另外，还有些配体就是机体的内源性分子，如热休克蛋白（heat shock protein，Hsp）60、HSP70、高迁移率族蛋白 B1（high mobility group box 1，HMGB1）等。表 13-1 列出了人类 11 个 TLR 成员的基因定位、存在部位及其主要配体。TLRs 与其配体结合后便可启动炎症反应。

表 13-1　人类 TLR 的种类、基因定位、存在部位及其主要配体

TLR 种类	基因定位	存在部位	主要配体
TLR1	4p14	广泛分布于各类免疫细胞膜表面	孔蛋白（porins）、肽聚糖、糖脂和非典型脂多糖等
TLR2	4q32	除 T、B、自然杀伤细胞外的免疫细胞膜表面，成纤维细胞、星形胶质细胞表面	多种 PAMPs（如肽聚糖、脂磷壁酸、脂蛋白等）以及内源性配体 HMGB1
TLR3	4q35	特异性分布于树突状细胞内	病毒双链 RNA、Poly（I：C）
TLR4	9q32-33	除 T、B、自然杀伤细胞外的免疫细胞表面，人胚肾细胞、心肌细胞、微血管内皮细胞、气道上皮细胞、脂肪细胞和肠上皮细胞等膜表面	LPS 及内源性配体 HSP60、HSP70、HMGB1 等
TLR5	1q41-42	除 T、B、自然杀伤细胞外的免疫细胞膜表面	细菌鞭毛蛋白
TLR6	4p14	广泛表达于多种细胞膜表面	脂磷壁酸、肽聚糖、脂肽
TLR7	Xp22.3	广泛表达于多种细胞内	单链 RNA，合成的核酸类似物等
TLR8	Xp22	同 TLR7	同 TLR7
TLR9	3p21.3	同 TLR7	细菌非甲基化 CpG DNA
TLR10	4p14	淋巴样组织和脾脏细胞表面	不直接识别配体，而是与其他 TLR 形成异二聚体发挥作用
TLR11	不详	巨噬细胞及肝、肾和膀胱上皮细胞膜表面	Profilin 样蛋白分子

TLR8 也有小部分表达于细胞表面。TLR2 能与 TLR1 或 TLR6 形成异源二聚体；TLR3、TLR4 和 TLR9 可形成同源二聚体；TLR5 可形成同源二聚体或与 TLR4 形成异源二聚体；TLR10 可形成同源二聚体或与 TLR1 形成异源二聚体，或与 TLR1/2 形成复合物

二、理化伤害性刺激信号受体

理化伤害性刺激导致炎症反应的机制主要有两种。第一种是直接损伤细胞，例如：①深度低温可使细胞内容物结晶，并最终损伤细胞膜，造成细胞内容物外泄；高温条件下，细胞膜通透性增加；而极高温则会导致血管的凝固和阻塞及生物大分子的变性，并引起细胞损伤；②一氧化碳、二硫化物、甲醛等化学物质也可引起细胞损伤，如一氧化碳可以通过与氧竞争性结合血色素而造成机体缺氧性损伤。损伤的细胞将引发机体清除受损细胞的炎症反应。第二种机制是通过受体介导引发的细胞功能改变。虽然有些理化因素不足

以直接损伤细胞，但是可以通过细胞表面的伤害性刺激受体介导细胞功能改变。这些受体均为细胞表面的蛋白分子，如压力受体、温度受体等。它们在受到伤害性刺激后，可以向细胞内转导这些伤害性刺激的信号，并通过外周和中枢神经系统诱导机体对这些刺激产生保护性反应。例如：酸敏感离子通道受体（Acid sensing ionic channel，ASIC）、瞬时电位受体（transient receptor potential，TRP）成员如辣椒素受体 1（vanilloid receptor 1，VR1）、辣椒素受体样受体 1（VR-like receptor 1，VRL-1）、EnaC/DEG 通道受体（epithelial Na^+ channel/degenerin）等均属此类内传理化伤害性刺激信号的受体（图 13-1）。

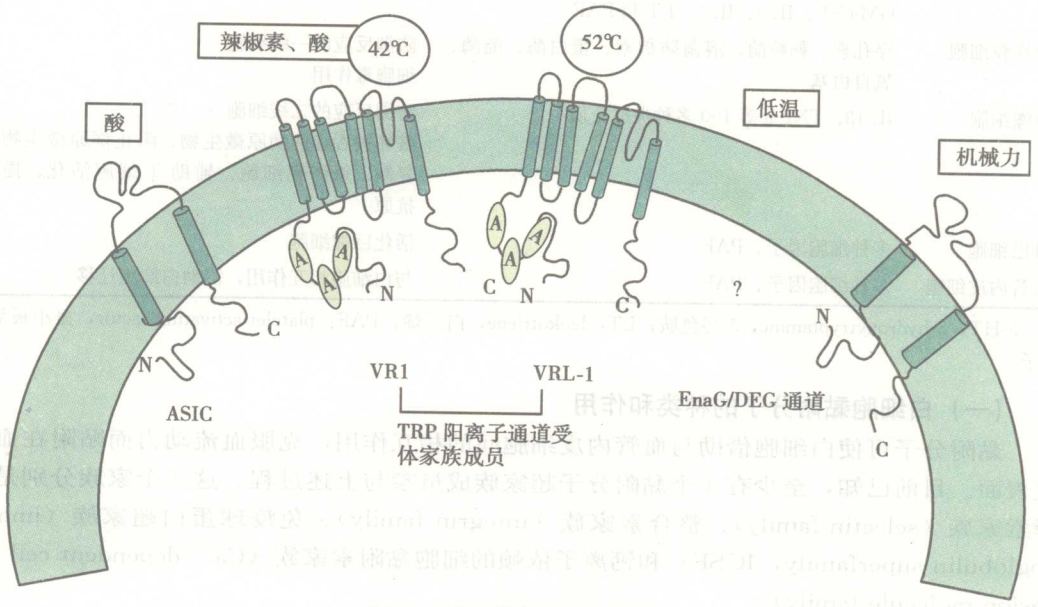


图 13-1　伤害性刺激受体示意图

ASIC：Acid Sensing Ion Channel，酸敏感离子通道；VR1：Vanilloid Receptor 1；VRL-1：Vanilloid Receptor like 1；TRP：transient receptor potential cation channel；EnaG/DEG：epithelial sodium channel/ degenerin

第二节　参与炎症反应的细胞及其作用机制

炎症发生的主要目的是将液体、蛋白质和细胞从血液输送至损伤组织，以清除病原、修复损伤。本节将主要讨论参与炎症反应的细胞及其被募集到损伤组织的分子机制。

一、参与炎症反应的细胞

参与炎症反应的主要细胞及其主要作用见表 13-2。

二、炎症反应细胞聚集的分子机制

血循环中的白细胞从血流聚集到炎症组织主要经过 3 个步骤：①克服血流动力而黏附在血管内皮细胞的表面，并沿血管壁爬行滚动，这一运动主要是通过白细胞黏附分子与其受体的相互结合来介导；②中性粒细胞和内皮细胞由于黏附分子与其受体的高亲和力相互作用以及细胞因子和趋化因子提供的刺激信号而活化；③趋化因子释放而形成的浓度梯度

诱导白细胞通过内皮细胞之间的缝隙穿过血管壁基底膜，进入组织间隙。

表 13-2　参与炎症反应的细胞及其主要作用

细胞种类	释放的炎性介质	主要作用
肥大细胞	组胺、5-HT、蛋白聚糖、趋化因子	小动脉扩张、小静脉收缩、血管通透性增加、致痛、刺激其他白细胞活化
嗜碱性粒细胞	同上	同上
嗜酸性粒细胞	IL-1β、IL-6、IL-8、TNF-α、TGF-α、TGF-β、GM-CSF、IL-3、IL-5、LT 和 PAF	刺激其他白细胞活化
中性粒细胞	穿孔素、颗粒酶、溶菌防御素、蛋白酶、酯酶、氧自由基	防御反应的一线细胞 细胞毒作用
巨噬细胞	IL-1β、TNF-α 等 100 多种炎症介质	防御反应的二线细胞 破坏被吞噬的病原微生物、阻止病原微生物的复制、杀伤靶细胞、辅助 T 细胞活化、提呈抗原
淋巴细胞	多种细胞因子、PAF	活化巨噬细胞
血管内皮细胞	多种细胞因子、PAF	与白细胞相互作用，协助白细胞迁移

5-HT，5-hydroxytryptamine，5-羟色胺；LT：leukotriene，白三烯；PAF：platelet-activating factor，血小板活化因子

（一）白细胞黏附分子的种类和作用

黏附分子可使白细胞借助与血管内皮细胞间的相互作用，克服血流动力而贴附在血管壁表面。目前已知，至少有 4 个黏附分子超家族成员参与上述过程。这 4 个家族分别是选择素家族（selectin family）、整合素家族（integrin family）、免疫球蛋白超家族（immunoglobulin superfamily，IGSF）和钙离子依赖的细胞黏附素家族（Ca^{2+}-dependent cell adhesion molecule family）。

选择素家族主要包括三个成员，分别是：L-selectin（CD62L）、E-selectin（CD62E）和 P-selectin（CD62P）。L、P 和 E 分别表示这三种选择素最初发现于白细胞、血小板和血管内皮细胞。

整合素是一大类异源二聚体糖蛋白，依据其中 β 亚单位的不同，又可分为一些亚家族。表达于白细胞表面的整合素不同于其他广泛存在的整合素，被称为白细胞整合素。白细胞整合素主要有 3 种异源二聚体分子，每种二聚体分子都含有相同的 95 kDa 的 β 亚单位（即 CD18），而 α 亚单位则分别为 CD11a（180 kDa）、CD11b（165 kDa）和 CD11c（150 kDa）。这 3 种白细胞整合素分别称为淋巴细胞功能相关抗原 1（lymphocyte function-associated antigen-1，LFA-1，即 CD11a/CD18）、补体受体 3（complement receptor type 3，CR3，即 CD11b/CD18）和 CR4（即 CD11c/CD18）。LFA-1 主要存在于淋巴细胞、单核细胞、巨噬细胞、粒细胞等，其配体是细胞间黏附分子 1（intracellular adhesion molecule-1，ICAM-1，即 CD54）和 ICAM-2（即 CD102）；CR3 和 CR4 主要在髓系细胞表达，可以与补体片段结合，从而协助吞噬细胞吞噬已被补体包被的颗粒。

属于 IGSF 的黏附分子有 ICAM-1、ICAM-2、血小板-内皮细胞黏附因子 1（platelet endothelial adhesion molecule-1，PECAM-1/CD31）等。白细胞穿过血管壁的过程与这些黏附分子的作用密切相关。ICAM-1 和 ICAM-2 主要表达于血管内皮细胞，PECAM-1 主要表达于血小板、血管内皮细胞、中性粒细胞、单核细胞。

参与白细胞和血管内皮细胞相互作用的主要黏附分子见表 13-3。白细胞黏附分子的缺陷将导致炎症反应不足，患者常出现反复感染，甚至致命性感染。

表 13-3　主要黏附分子的种类和作用

	中性粒细胞黏附分子	血管内皮细胞相应的黏附分子	作　用
选择素相互作用	p150sLeX（CD15） sLeX（CD15） L-选择素（CD62L）	P-选择素（CD62P） E-选择素（CD62E） GlyCAM	白细胞移动及其与血管内皮细胞的结合
整合素-IGSF 相互作用	LFA-1（CD11a/ CD18） VLA-4（CD49d/CD29）	ICAM-1（CD54） ICAM-2（CD102） VCAM-1（CD106）	继发性黏附、伸展、在炎症组织的归巢
IGSF 相互作用	PECAM-1（CD31） HCAM（CD44）	PECAM-1（CD31）	增强黏附、白细胞跨血管壁基质的移动、与透明质酸及其他细胞外基质结合

GlyCAM：Glycosaminoglycan-cell adhesion molecule，糖胺聚糖细胞黏附分子；HCAM：homing cell adhesion molecule，归巢细胞黏附分子；sLeX：sialylated Lewis antigen X，唾液酸化的路易斯寡糖 X；VCAM-1：vascular cell adhesion molecule 1，血管细胞黏附分子 1；VLA-4：very late antigen 4；迟现抗原 4

（二）白细胞迁移和趋化的分子机制

白细胞向炎症部位的迁移是在趋化物精确调控下发生的定向移动。这种细胞随着环境中化学物质分布的梯度进行的定向迁移称为趋化性迁移（chemotaxis）。因趋化物包括吸引和排斥细胞的物质，故细胞的运动可以朝向趋化物所在部位，也可以朝向趋化物所在部位的相反方向。

趋化物有内源性和外源性之分。外源性趋化因子主要为细菌中类似 FMLP（N-formyl-methionyl-leucyl-phenylalanine）的寡肽、凝集素、变性蛋白、某些脂类和脂多糖等。

内源性趋化物由宿主细胞产生，其中最重要的一大类分子是具有趋化作用的细胞因子，简称趋化因子（chemokines）。趋化因子是趋化性迁移的关键调节者，在炎症反应的启动和进展过程中具有至关重要的作用。趋化因子的分子量在 8 kDa 到 11 kDa 之间，作用浓度在 1 ng/ml 到 100 ng/ml，由多种细胞类型所产生。趋化因子的主要作用有：①聚集在内皮细胞表面，当白细胞在选择素介导下沿着血管内皮滚动时，作为白细胞停留的特异信号；②可选择性激活白细胞，活化的白细胞停止滚动，相对紧密地贴附在内皮；③在血管内皮表面形成浓度梯度，引导黏附的白细胞向着高浓度趋化因子方向移动；④使白细胞渗出血管，迁移到组织间隙，在血管外，仍然朝高浓度趋化因子方向移动。

依据结构特点，可将趋化因子分为四类，即“C-X-C”类（在两个半胱氨酸残基之间间隔一个其他氨基酸残基，X 代表任何氨基酸残基）、“C-C”类（两个半胱氨酸残基相连）、“C”类（独立的半胱氨酸残基）及其他类。主要趋化因子及其在炎症中的作用见表 13-4。各种趋化物在相应的被趋化细胞表面均有受体存在，配体与受体的相互作用使细胞内多种分子发生一系列生物物理和生物化学改变，其具体信号转导机制见本章第四节。

表 13-4　代表性内源趋化因子的种类和作用

趋化因子的种类	作用细胞					
	中性粒	单核/巨噬	嗜酸	嗜碱	淋巴	NK
C-X-C 类趋化因子						
IL-8、NAP-2	+	−	−	−	±	−
C-C 类趋化因子						
MCP-1	−	+	−	+	+	+
MCP-3、MIP-1、RANTES	−	+	+	+	+	+

续表

趋化因子的种类	作用细胞					
	中性粒	单核/巨噬	嗜酸	嗜碱	淋巴	NK
C类趋化因子						
淋巴细胞趋化素	−	−	−	−	+	+
其他类趋化因子						
C5a	+		+	−	−	−
LTB4、PAF	+	+	−	−	−	−
PDGF	+	+	−	−	−	−
IL-3、IL-5	−	−	+	+	−	−
TGF-β	+	±	−	−	+	+

NAP-2：neutrophil-activating protein 2，中性粒细胞活化蛋白 2；MCP-1：monocyte chemoattractant protein 1，单核细胞趋化蛋白 1；MIP：macrophage inflammatory protein，巨噬细胞促炎蛋白；RANTES：regulated upon activation normal T expressed and presumably secreted chemokine；LTB4：leukotrienes B4，白三烯 B4

第三节　主要炎症介质及其作用机制

炎症介质（inflammatory mediator）是介导炎症反应的化学分子的统称，属于可溶扩散性分子，因此能在感染或组织损伤的局部或扩散至远处发挥作用。炎症介质可来自细胞和血浆，来自细胞的炎症介质可在细胞内以颗粒形式贮存，在需要时释放至胞外，也可在某些致炎因子的刺激下即刻合成；来自血浆的炎症介质通常以前体形式存在，经蛋白酶水解而激活。多数炎症介质通过与靶细胞表面的受体结合发挥其作用，也有些炎症介质直接有酶活性或者可介导氧化损伤。

一、细胞释放的炎症介质

（一）细胞因子

参与炎症反应的细胞因子（cytokine）有促炎和抗炎两大类。促炎细胞因子主要由活化的巨噬细胞产生，能促进各种炎症病理过程的发生，如白细胞介素（interleukin，IL）-1β、IL-6、IL-11、肿瘤坏死因子（tumor necrosis factor，TNF）-α、干扰素（interferon，IFN）-α、IFN-β 以及趋化因子等，其中最重要的是 IL-1β 和 TNF-α，二者的拮抗剂有的已用作临床抗炎药物。

抗炎细胞因子主要由 T 细胞产生，可抑制炎症反应的进一步发展，如 IL-4、IL-10、IL-13 等。TGF-β 可抑制促炎细胞因子 IL-1β、IL-6 和 TNF-α 的生成，因此可认为属抗炎细胞因子。不过，TGF-β 本身亦有促炎细胞因子的活性，包括对 T 淋巴细胞和中性粒细胞的趋化作用等。因此，TGF-β 在体内兼具抗炎和促炎双重作用。抗炎细胞因子在临床上可用于某些炎性疾病的治疗。代表性的促炎和抗炎细胞因子见表 13-5。

需要强调的是，炎症反应中，在促炎细胞因子、释放促炎细胞因子的细胞和促炎细胞因子的靶细胞之间存在着一个极其复杂的调控网络（图 13-2）。

（二）血管活性胺

参与炎症反应的血管活性胺主要有组胺和 5-HT。组胺是最早发现的一种炎症介质，由组氨酸脱羧后生成。组胺生成后储存于肥大细胞和嗜碱性粒细胞的颗粒中。引起组胺释放的因素很多，主要是 IgE 介导的反应，某些理化因素也可诱导组胺的释放。组胺在炎症反应中是最重要的血管活性介质，其作用包括血管扩张、血管内皮细胞收缩引起的血管

表 13-5　主要促炎和抗炎细胞因子的种类和作用

作　用	细胞因子
促炎细胞因子	
致热源	IL-1β、TNF-α、IL-6
放大炎症反应	IL-1β、TNF-α、IL-6、IFN-α、IFN-β、趋化因子
促进急性时相反应物的产生	IL-1β、IL-6、IL-11、TNF-α、IFN-γ、TGF-β、LIF、OSM、CNTF
趋化因子	
C-X-C 类趋化因子	IL-8、PF-4、PBP、NAP-2、β-TG
C-C 类趋化因子	MIP-1α、MIP-1β、MCP-1、MCP-2、MCP-3、RANTES
C 类趋化因子	lymphotactin
促进促炎细胞因子产生	IL-12
抗炎细胞因子	
抑制促炎细胞因子产生	IL-4、IL-10、IL-13、TGF-β

LIF：leukemia inhibitory factor，白血病抑制因子；OSM：oncostatin M，抑癌蛋白 M；CNTF：ciliary neurotrophic factor，睫状神经趋向因子；PF-4：platelet factor 4，血小板因子 4；PBP：platelet basic protein，血小板碱性蛋白；β-TG：β-thromboglobulin，β-血栓球蛋白

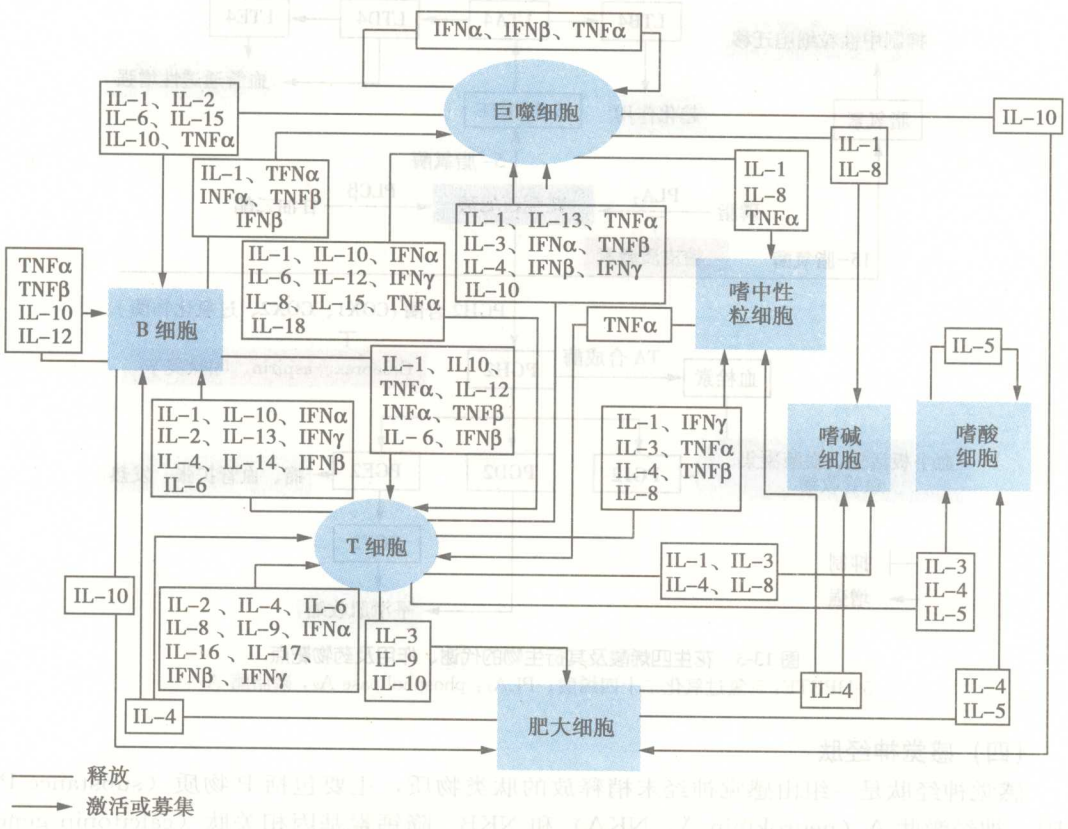


图 13-2　炎症反应中的细胞因子网络

通透性增高、非血管平滑肌收缩、募集嗜酸性粒细胞、阻断 T 淋巴细胞功能等。组胺在细胞表面有 H1、H2 和 H3 三种受体，分别介导不同反应。H1 受体主要介导急性血管反应、支气管平滑肌收缩、吸引嗜酸性粒细胞；H2 受体具有抑制嗜酸性粒细胞分化的抗炎效果，但是仍可以导致血管扩张；H3 受体则主要抑制各种细胞合成和释放组胺，使组胺的致炎作用减弱或消失；这是机体维持生理平衡需要的一种自限方式。

5-HT 的作用与组胺类似，其主要存在于血小板和内皮细胞。胶原纤维、凝血酶、ADP、免疫复合物、PAF 等可促进 5-HT 从细胞释放。

（三）花生四烯酸及其衍生物

花生四烯酸是炎症介质前列腺素（prostaglandin，PG）和白三烯（LT）的重要中间产物。在环氧化酶（cyclooxygenase，Cox）的作用下，花生四烯酸生成环状结构的 PGs；在脂氧化酶作用下，花生四烯酸生成线状结构的 LT 系列衍生物。这一过程的相关代谢酶可作为抗炎药物的作用靶点（图 13-3）。例如：阿司匹林和非甾体类抗炎药物可抑制 Cox 活性，抑制 PG 的产生。糖皮质激素可抑制 Cox-2 的活性，抑制 IL-1β 和 TNF-α 的表达，并能上调某些抗炎蛋白的表达，发挥抗炎作用。PG 类物质在炎症中的主要作用有：①舒张血管；②致支气管、胃肠和子宫平滑肌收缩；③促进腺体分泌；④作为趋化因子；⑤致发热。LT 类脂氧化酶的产物主要有 LTB4、LTC4、LTD4 和 LTE4。LTB4 是中性粒细胞的趋化因子和白细胞功能反应（黏附于内皮细胞、产生氧自由基和释放溶酶体酶）的激活因子；LTC4、LTD4 和 LTE4 可致血管收缩、支气管痉挛和静脉血管通透性增加。

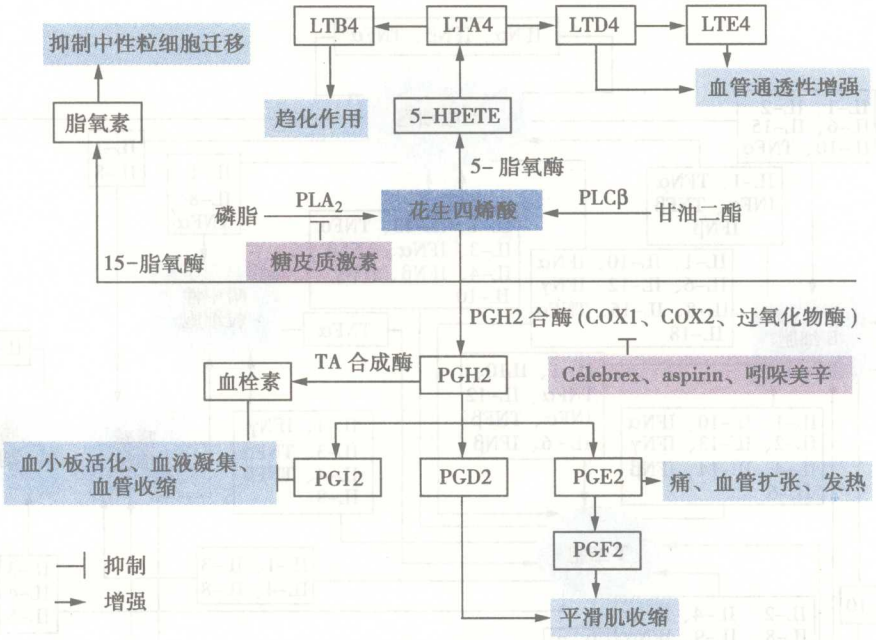


图 13-3　花生四烯酸及其衍生物的代谢、作用及药物靶点

5-HPETE：5-氢过氧化二十四烯酸；PLA_2：phospholipase A_2，磷脂酶 A_2

（四）感觉神经肽

感觉神经肽是一组由感觉神经末梢释放的肽类物质，主要包括 P 物质（substance P，SP）、神经激肽 A（neurokinin A，NKA）和 NKB、降钙素基因相关肽（calcitonin gene-related peptide，CGRP）等。

感觉神经肽类物质具有明显的促炎作用，能通过轴突反射机制引起神经源性炎症及加重炎症反应（图 13-4）。CGRP 对人的皮肤是一种很强的血管扩张剂，大剂量时可引起血管通透性增加，并形成荨麻疹。SP 能引起血管扩张外，更是一种极强的致水肿因子。

此外，活化的中性粒细胞和单核细胞所释放的氧自由基和溶酶体酶，也可作为炎症介质，促进炎症反应和破坏组织；多种炎细胞产生的磷脂类介质血小板活化因子（PAF）可

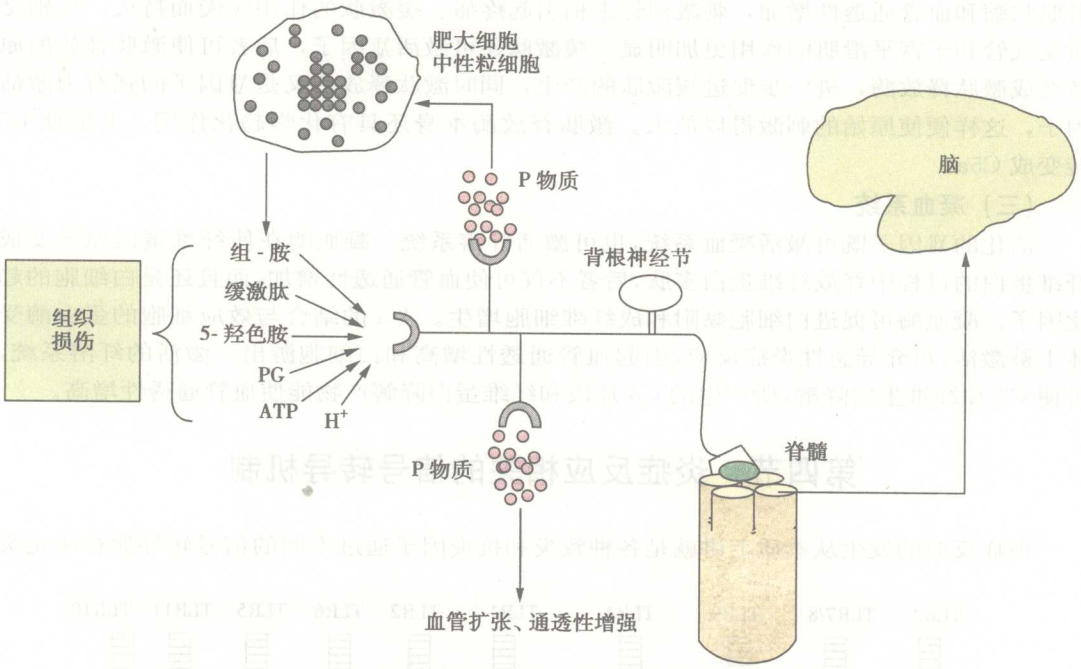


图 13-4　感觉神经肽参与神经系统对炎症反应的影响
在致炎因子的作用下，从感觉神经末梢释放的 SP，能使邻近的肥大细胞释放组胺，而组胺和激肽等炎症介质又可刺激感觉神经末梢释放 SP

凝聚并活化血小板，使其释放活性胺类，引起毛细血管扩张和通透性增加并募集嗜酸性粒细胞；一氧化氮（NO）也是一种抗炎分子，可引起小血管扩张，抑制血小板的黏附、激活、聚集和脱颗粒，抑制肥大细胞引起的炎症反应，抑制白细胞向炎症灶集中。

二、体液中的炎症介质

（一）补体系统

补体系统由 30 种以上的糖蛋白组成，分别存在于血清或细胞膜表面。血流中的补体成分以非活化形式存在，可通过三种方式被激活，即免疫复合物诱发的经典激活途径，糖、凝血素诱发的凝集素途径和菌体产物诱发的旁路激活途径。补体系统在炎症中的主要作用见表 13-6。

表 13-6　补体系统在炎症中的主要作用

补体组分和片段	主要作用
C5a、C3a、$C5a^{desArg}$、C4a	增加血管通透性，促平滑肌收缩，促肥大细胞和嗜碱性粒细胞脱颗粒
C5a、$C5a^{desArg}$	活化及趋化中性粒细胞，促进 PG 和 LT 合成
C3b、C4b	促进吞噬细胞吞噬细菌和免疫复合物
C3b、C5a、$C5a^{desArg}$、C1q	活化专职吞噬细胞
C5b	溶解外源细菌或细胞
C3b、CR1	溶解循环中的免疫复合物

（二）激肽系统

激肽系统是血液中除补体外的第二大炎症介质形成系统，其多种中间产物与补体系统互有联系。激肽系统激活的最终产物主要是缓激肽（由 9 个氨基酸残基组成），其可致平

滑肌收缩和血管通透性增加，刺激神经末梢引起疼痛。缓激肽的作用缓慢而持久，特别是对支气管和子宫平滑肌的作用更加明显。缓激肽还可激活Ⅻ因子，后者可使激肽释放酶原转变成激肽释放酶，进一步促进缓激肽的产生，同时激肽释放酶又是Ⅻ因子的强有力激活因子，这样便使原始的刺激得以放大。激肽释放酶本身还具有化学趋化作用，并能使 C5 转变成 C5a。

（三）凝血系统

活化的Ⅻ因子既可激活凝血系统，也可激活纤溶系统。凝血酶在使纤维蛋白原转变成纤维蛋白的过程中释放纤维蛋白多肽，后者不仅可使血管通透性增加，而且还是白细胞的趋化因子。凝血酶可促进白细胞黏附和成纤维细胞增生。Ⅹa 的结合与效应细胞的蛋白酶受体Ⅰ被激活，可介导急性炎症反应，引起血管通透性增高和白细胞游出。激活的纤溶系统，可使 C3 和纤维蛋白降解，所产生的 C3 片段和纤维蛋白降解产物能使血管通透性增高。

第四节　炎症反应相关的信号转导机制

炎症反应的发生从本质上讲就是各种致炎和抗炎因子通过不同的信号转导途径决定炎

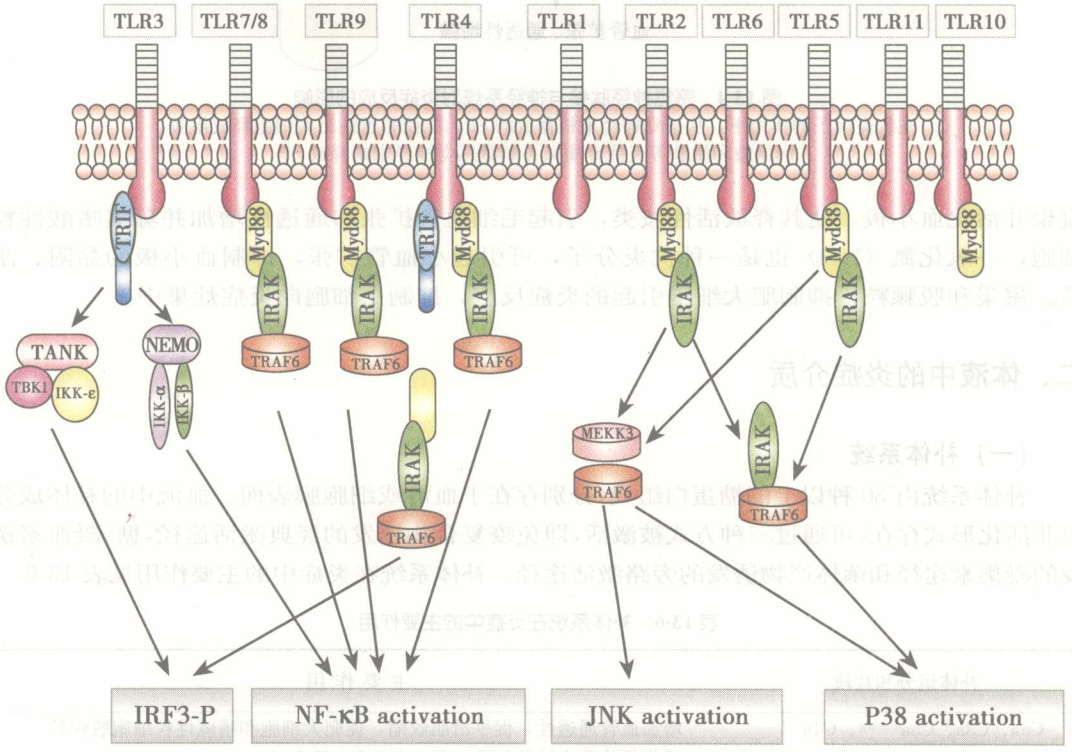


图 13-5　TLRs 介导的信号转导通路

图中的膜结构指细胞质膜或位于胞浆中的内涵体（endosome）膜。胞浆中含有 TIR 结构域的接头分子主要有 5 种，包括髓样分化因子 88（myeloid differentiation factor 88，MyD88）、MyD88 接头样分子（MyD88-adaptor like，Mal），又称 TIRAP（TIR domain-containing adaptor protein）、含有 TIR 结构域的干扰素 β 诱导接头蛋白（TIR domain-containing adaptor inducing interferon β，TRIF）、TRIF 相关的接头分子（TRIF-related adaptor molecule，TRAM）以及含有 sterile α 和 armadillo 模体的蛋白（sterile α-and armadillo-motif-containing protein，SARM）。除 TLR3 外，所有 TLRs 均可募集 MyD88。只有某几个 TLRs 可募集 Mal、TRAM 和 TRIF，以显示在信号转导过程中的特殊性。一种 TLR 可激活多条信号转导通路，有的通路被多个 TLRs 所共享，而有的通路则是某些 TLRs 特有的。导致不同通路活化的原因在很大程度上是由 TLRs 的胞内结构域所募集的不同接头分子决定的

症反应细胞的生物学效应的过程。在炎症反应过程中，不同的炎症刺激可引起细胞内不同信号途径的激活；而对于不同细胞，同一种刺激也可能会引起不同信号途径的激活。由于细胞内信号途径之间的作用不是孤立的，细胞的多种生物学功能都是通过多条信号途径的网络来动态调节的。各种受体介导的信号转导的共同机制已在本书第八章中有详细讨论，本节仅介绍两条与炎症反应密切相关的信号转导通路。

（一）TLRs 介导的信号转导

TLRs 介导信号转导的基本过程如图 13-5 所示：①配体与位于细胞膜或位于细胞内的 TLRs 结合，使受体活化；②活化的 TLRs 募集胞浆中含有 TIR 结构域的接头分子，使接头分子活化；③通过上述接头分子进一步募集和活化下游信号转导分子，经过多个步骤最终激活 NF-κB 和/或 P38、JNK 或干扰素调节因子 3（interferon regulatory factor 3，IRF3）。NF-κB 和 IRF3 均为转录因子，二者被活化后，可通过与 DNA 顺式作用元件结合而调节相关基因表达；P38 和 JNK 均为 MAPK 家族的蛋白激酶，二者被活化后，可使转录因子磷酸化，进而调节有关基因的转录。另外，图 13-6 以 TLR4 介导 LPS 信号的转导过程为例，较具体地描述了 TLRs 介导的信号通路。

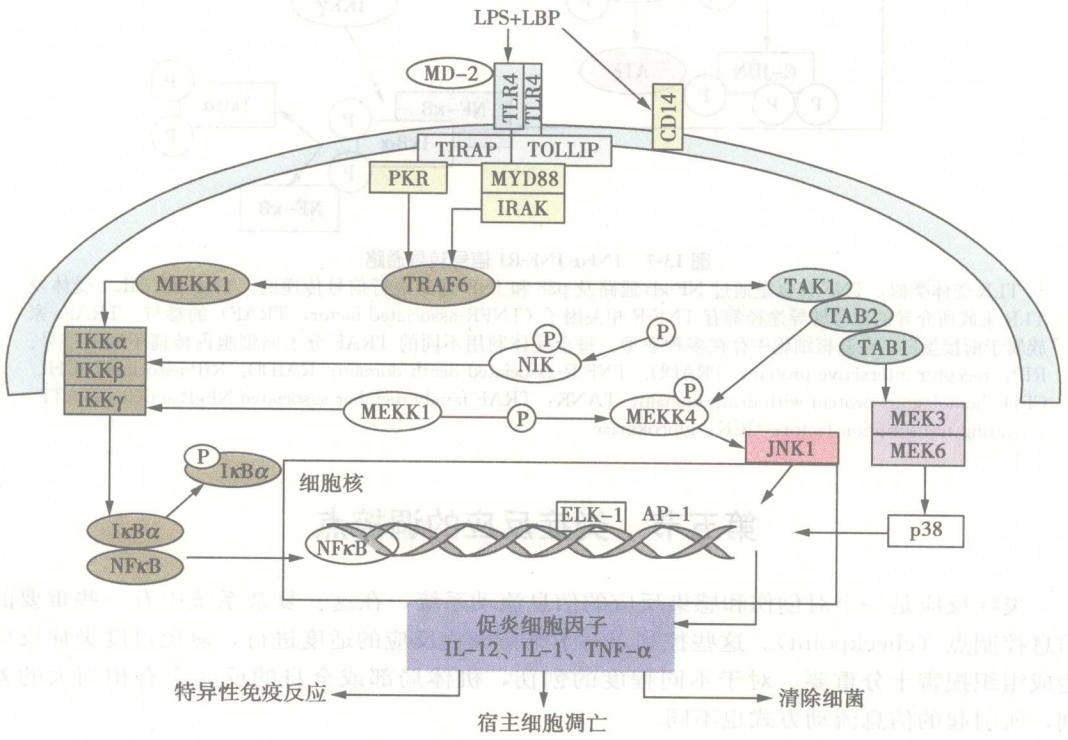


图 13-6　TLR4 受体介导的信号转导及细胞效应

TLR4 介导 LPS 信号的转导过程：①LBP 携带 LPS 结合至 TLR4，CD14 辅助这一结合；②活化的 TLR4 募集并活化胞浆中含有 TIR 结构域的接头分子；③通过上述接头分子进一步募集和活化下游各种衔接蛋白和蛋白激酶，最后分别激活 NF-κB 通路、P38 介导的通路及 JNK 介导的通路

（二）TNF-α 受体介导的 TNF-α 信号转导

TNF-α 与相应受体 TNF 受体 1（TNFR1）结合后的主要信号转导通路如图 13-7 所示。

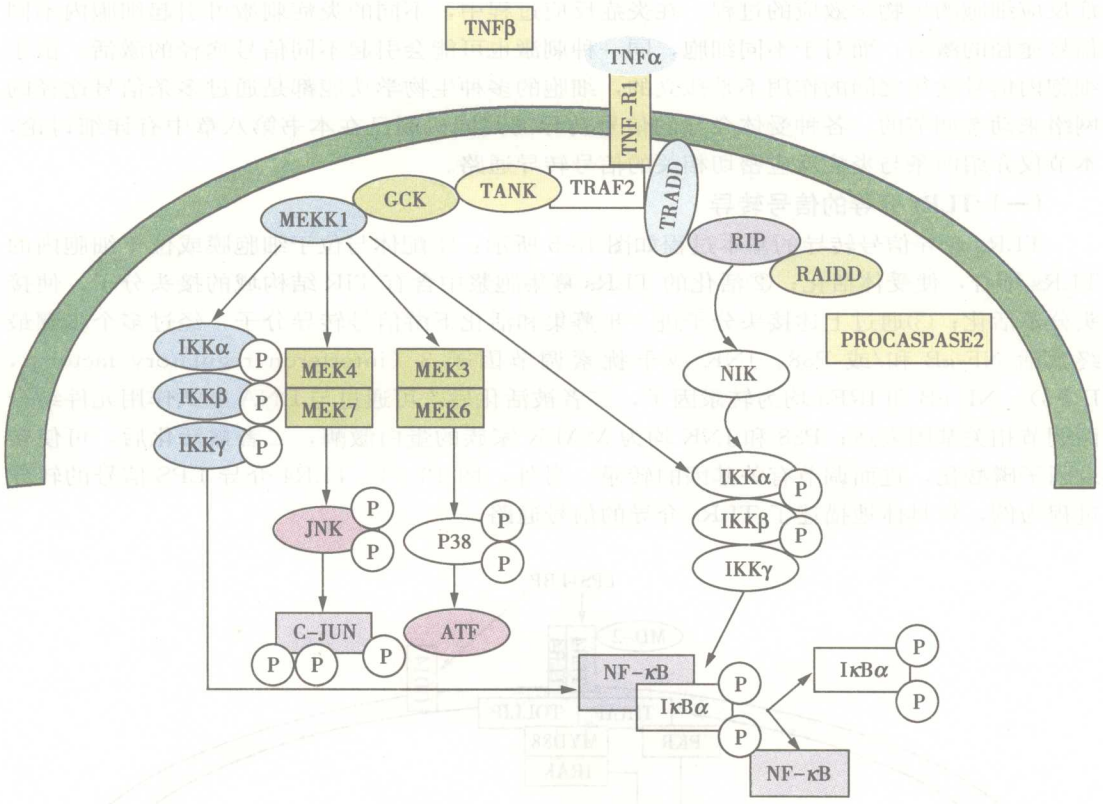


图 13-7　TNFα-TNF-R1 信号转导通路

与 TLR 受体类似，TNF-R 也是通过 NF-κB 通路及 p38 和 JNK 通路进行信号传递的。TNF-R1、IL-1 受体及 TLR 家族所介导的信号转导途径都有 TNF-R 相关因子（TNFR-associated factor，TRAF）的参与。TRAFs 家族属于衔接蛋白，在真核细胞中存在多种亚型。每个受体利用不同的 TRAF 分子向细胞内传递不同的信号。RIP：receptor interactive protein；TRADD：TNF-R-associated death domain；RAIDD：RIP-associated ICH1/CED3-homologous protein with death domain；TANK：TRAF family member-associated NFκB activitor；ATF：activating transcription factor；GCK：glucokinase

第五节　炎症反应的调控点

炎症反应是一个对创伤和感染反应的信息流动系统，在这一复杂系统中有一些重要的信息控制点（checkpoint）。这些控制点对于维持炎症反应的适度进行，避免过度炎症反应造成组织损害十分重要。对于不同程度的创伤，机体局部或全身的反应会有相当大的差别，所引起的信息流动方式也不同。

一、早期炎症正向信号

早期炎症反应由位于组织中的肥大细胞和巨噬细胞的活化所启动。创伤带来局部神经肽、HSP、线粒体中的甲酰化肽等刺激肥大细胞和巨噬细胞，使其释放出大量炎症介质，这些炎症介质可激活血流中的中性粒细胞、淋巴细胞等，并使其外渗迁移至损伤及感染部位。图 13-8 中的箭头表示细胞可以被激活和募集。

尽管机体对于创伤和感染的早期反应几乎是无延搁的加速反应过程，但是随后发生的细胞反应则是有节制的。具体表现在，除了肥大细胞和巨噬细胞可以直接被组织损伤或感

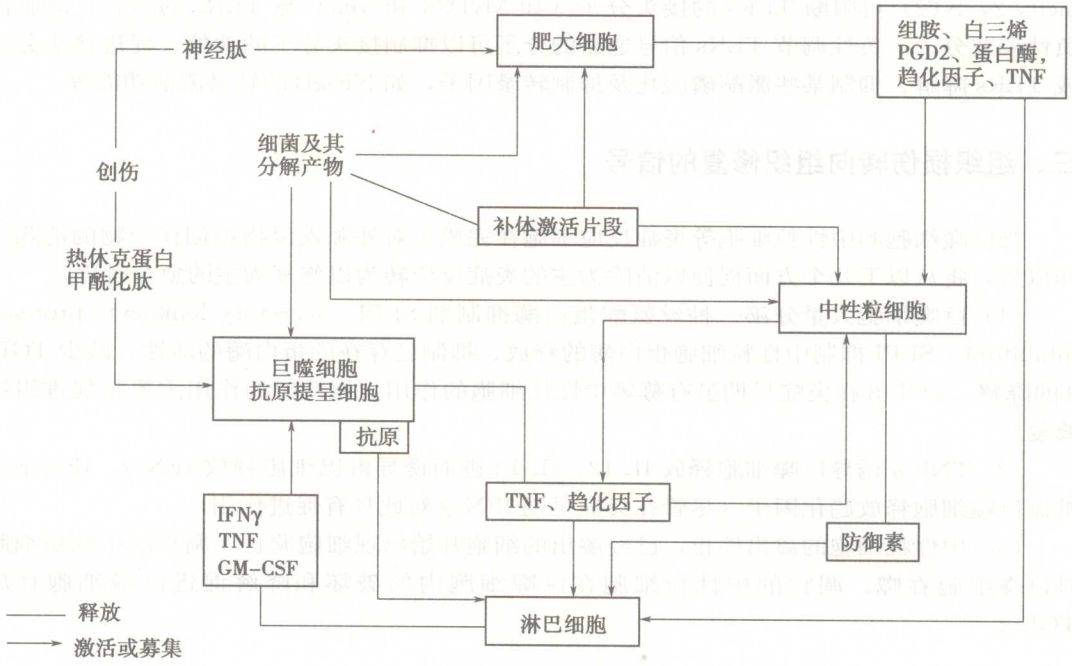


图 13-8　早期炎症正向信号

染细菌的产物所活化外，其他炎细胞的活化，包括巨噬细胞的完全活化都是受双信号系统控制的。淋巴细胞、中性粒细胞的活化以及巨噬细胞的完全活化除了需要各种炎症介质作为第一刺激信号，还需要来自原发损伤组织和细菌产物的第二信号。这种双信号控制，保证了炎症反应发生的准确性。

二、早期炎症终止信号

炎症反应可受到反馈调节。促使早期炎症反应终止的信号主要有以下 4 种。

1. 脂氧素　脂氧素（lipoxin，LX）为花生四烯酸脂氧合酶的代谢产物，是一类重要的内源性脂质抗炎介质，因其具有强力的抗炎效应而被称为炎症反应的“刹车信号”。其抗炎作用主要表现在：①通过下调趋化因子和黏附分子的表达而抑制中性粒细胞趋化、黏附和渗出，显著减轻炎症部位中性粒细胞的浸润；另外还能抑制炎性刺激引起的中性粒细胞产生活性氧和炎性细胞因子的表达；②促进巨噬细胞对凋亡中性粒细胞的吞噬，有效地避免凋亡中性粒细胞发生继发性坏死和致炎，从而间接发挥抗炎作用；③抑制促炎细胞因子（如 TNF-α）的表达，同时促进抗炎细胞因子（如 TGF-β）的表达。

2. 腺苷　细胞膜表面的 CD39 分子可促进从活化的炎症反应细胞中释放的、或从破损细胞漏出的 ATP 和 ADP 的水解，产生的腺苷通过邻近细胞表面的腺苷受体抑制这些细胞的炎症反应。

3. 胆碱能神经递质　自主神经系统的胆碱能神经递质可阻止巨噬细胞释放 TNF-α。

4. TLRs 信号通路中的负性调节分子　TLRs 信号通路中的负性调节分子很多，按存在部位可分为：①胞外负性调节分子。可溶性 TLRs 是最重要的胞外负性调节分子，通过竞争结合 TLRs 的激动剂而有效削弱 TLRs 的正向信号，从而阻断急性炎症反应；②胞膜上的负性调节分子。通过阻断 TLRs 接头分子与 TLRs 胞内结构域的结合或干扰配体与特异性 TLRs 的结合而负性调节炎症反应。如 肿瘤生成抑制子 2（suppressor of tumorige-

nicity 2，ST2）可阻断 TLRs 的接头分子（如 MyD88 和 Mal）与 TLRs 的结合；③胞内负性调节分子。负性调节 TLRs 信号通路的分子可以抑制接头分子的功能、促进接头分子或 TLRs 降解、抑制某些激酶磷酸化及抑制转录因子，如 NF-κB 的转录激活功能等。

三、组织损伤转向组织修复的信号

当巨噬细胞和中性粒细胞等炎症反应细胞在完成了对外源入侵物和创伤产物的清除工作以后，将从以下几个方面促使以清除为主的炎症反应转为以修复为主的愈合过程。

（1）巨噬细胞大量分泌一种丝氨酸蛋白酶抑制剂 SLPI（secretory leukocyte protease inhibitor）。SLPI 抑制中性粒细胞蛋白酶的释放、抑制已存在的蛋白酶的活性、减少 TGF-β 的降解。TGF-β 在炎症早期虽有募集中性粒细胞的作用，但此时的作用主要是促进组织修复。

（2）TNF-α 诱导巨噬细胞释放 IL-12。IL-12 进而诱导淋巴细胞释放 IFN-γ。IFN-γ 可抑制巨噬细胞释放趋化因子（尽管在炎症早期 IFN-γ 对此具有促进作用）。

（3）中性粒细胞的渗出停止，已经渗出的细胞开始出现细胞凋亡，凋亡的中性粒细胞被巨噬细胞吞噬。凋亡的中性粒细胞在巨噬细胞内的破坏和降解促进巨噬细胞释放 TGF-β。

四、中枢神经系统在炎症反应中的作用

中枢神经系统可反射性地实时调整机体的炎症反应。中枢神经系统在炎症中的调节作用主要体现在：①中枢神经元可合成和表达促炎细胞因子 TNF-α 和 IL-1 等。神经元也可对这些细胞因子产生反应。同时，这些细胞因子能兴奋下丘脑-垂体-肾上腺系统，促进糖皮质激素释放，而糖皮质激素可通过抑制细胞释放促炎介质而发挥抗炎作用。②副交感神经递质乙酰胆碱能有效地使巨噬细胞进入静止状态。炎症组织中的巨噬细胞表面有乙酰胆碱受体，结合乙酰胆碱后细胞 TNF-α 的释放明显降低。

（何凤田）

第十四章　心血管系统疾病的分子机制

关键知识点

- 心血管活性肽和对心血管功能具有调节效应的气体分子是控制心血管系统自稳态的物质基础，这些心血管活性物质平衡失调可引起高血压。
- 高血压和脂质代谢紊乱所造成的血管内皮损伤是动脉粥样硬化发病的始动环节，血管平滑肌细胞表型转化和异常增殖是动脉粥样硬化发生发展的主要细胞病理学事件，慢性炎症和氧化应激反应贯穿于动脉粥样硬化发生发展的全过程。
- 高血压和动脉粥样硬化引起的心脏负荷增加可通过促进生长因子和心血管活性肽的合成与分泌，诱导即刻早期基因、某些胚胎基因和心肌肥大相关基因表达而导致心肌肥大。
- 心律失常可继发于各种心血管疾病，也可直接由心肌细胞离子通道基因突变所引起，长 QT 综合征和房颤是最常见的家族性心律失常。

关键词：心血管活性肽；气体信号分子；高血压；动脉粥样硬化；心肌肥大；长 QT 综合征；心房颤动

对人类健康威胁最大的高血压、动脉粥样硬化等心血管疾病发病率和死亡率高、病因复杂，其分子机制目前尚不十分清楚。目前认为糖、脂和能量代谢紊乱引起的心血管活性物质平衡失调和血管内皮损伤可能是这一类疾病的始动环节。心血管疾病相关分子机制的阐明对于发现新的药治疗靶位至关重要。

第一节　心血管活性物质与高血压

心血管活性物质是指由心血管组织以旁分泌、自分泌以及内分泌方式分泌的对心血管功能具有调节效应的小分子生物活性物质，包括小分子活性多肽、氨基酸衍生物、脂类介质和气体信号分子等。这类小分子活性物质种类繁多，效应各异，相互作用形成非常复杂的调节网络，构成循环系统自稳态调节的物质基础，共同调节心血管系统的功能。

一、心血管活性肽分为缩血管活性肽和舒血管活性肽

心血管活性肽（vasoactive peptide）是指心血管活性物质中的一些短肽，所含氨基酸残基数少则几个，多则数十个。该类活性肽成员众多，功能各异。按功能不同，心血管活性肽可以分为两类：一类是收缩血管、促进血管平滑肌细胞增殖的活性肽；另一类是舒张血管、抑制血管平滑肌细胞增殖的活性肽。部分心血管活性肽的氨基酸残基数目、受体类型、生物学效应见表 14-1。

二、缩血管活性肽和舒血管活性肽之间的平衡失调引起高血压

人的高血压（high blood pressure）定义为在未服用抗高血压药物的情况下，收缩压大于或等于 18.7 kPa（140 mmHg）和（或）舒张压大于或等于 12 kPa（90 mmHg）。高

血压虽然是遗传因素与环境因素相互作用所诱发的多基因、多因素的复杂性疾病，但是缩血管活性肽和舒血管活性肽之间的平衡失调在高血压的发生和发展中起着重要作用。心血管活性肽除调节血管舒缩活动外，还可调节血管平滑肌细胞的增殖活性，因此，这类活性肽的平衡失调也是引起血管增殖性病变，即血管重塑（vascular remodeling）的主要原因。

表 14-1　一些重要的心血管活性肽及其生物学效应

心血管活性肽及其亚型	氨基酸残基数	受体类型	主要信号转导途径	生物学效应
内皮素（ET）	21	G 蛋白偶联受体，有主要存在于血管平滑肌的 ETA 和主要存在于血管内皮细胞的 ETB 两种	磷脂酰肌醇途径 MAPK 途径	收缩血管 促进血管平滑肌增殖 促进心肌细胞肥大
血管紧张素Ⅱ（AngⅡ）	8	G 蛋白偶联受体， 1 型受体：AT1R 2 型受体：AT2R	AT1R： MAPK 途径 JAK-STAT 途径 DG-PKC 途径 AT2R： cGMP-PKG 途径	AT1R： 促进血管和心脏收缩 促进血管平滑肌增殖 促进心肌细胞肥大 AT2R： 舒张血管 抑制血管平滑肌增殖 促进血管平滑肌分化
心钠素（ANP）	28	受体型鸟苷酸环化酶	cGMP-PKG 途径	舒张血管 抑制血管平滑肌增殖
降钙素基因相关肽(CGRP)	37	G 蛋白偶联受体	cAMP-PKA 途径	舒张血管 抑制血管平滑肌增殖
肾上腺髓质素（AdM）	52	G 蛋白偶联受体	cAMP-PKA 途径	舒张血管 抑制血管平滑肌增殖

（一）缩血管活性肽

内皮素-1（endothelin-1）和血管紧张素Ⅱ（angiotensin Ⅱ，Ang Ⅱ）在维持心血管功能、调节血压方面具有重要作用。高血压病患者常出现内皮素基因表达异常，肾素-血管紧张素系统处于激活状态。已经发现，内皮素-1 与 Ang Ⅱ之间存在正反馈调节机制。当血管内皮细胞功能受损、内皮素-1 合成与分泌增加时，内皮素-1 可促使血管平滑肌细胞释放 Ang Ⅱ，后者又进一步促进内皮素基因表达，结果导致血管强烈收缩。内皮素-1 不仅能强烈收缩冠状动脉，而且高浓度时还可直接损伤心肌细胞。临床上已将内皮素-1 水平作为评价高血压和动脉粥样硬化严重程度的一个指标。用受体阻断剂抑制内皮素-1 或 Ang Ⅱ与相应受体相互作用，或通过抑制血管紧张素转换酶（angiotensin converting enzyme，ACE）活性使 Ang Ⅱ生成减少，可起到治疗高血压病的作用，这种治疗策略已在临床上得到广泛应用。

（二）舒血管活性肽

1. 心钠素　心钠素（atrial natriuretic peptide，ANP）具有扩张血管、降低血压、抑制血管平滑肌细胞增殖和促进肾对水及钠的排泄作用，在高血压病患者和自发性高血压大鼠的血浆中水平明显升高。目前认为，这是机体的一种代偿及保护性反应。高血压病患者和实验性高血压动物血浆中 ANP 水平虽然较高，但由于 ANP 受体数量下调或受体对 ANP 的亲和力降低，以及体内存在 ANP 抗体，使机体对 ANP 的反应性降低，造成 ANP 相对不足。ANP 基因第 2 内含子的 Hpa II 和第 3 外显子的 Aca I 位点突变也参与高血压

病的发生和发展过程。促进内源性 ANP 合成和分泌，或静脉给予外源性 ANP 对高血压病具有显著的治疗效果。

2. 降钙素基因相关肽　降钙素基因相关肽（CGRP）对全身血管，包括冠状动脉，具有强大的舒张作用，是目前已知最强的舒血管物质。CGRP 通过两种机制舒张血管：一是血管内皮依赖的途径，作用于血管内皮细胞，通过促进舒血管活性物质的释放引起血管舒张；二是非内皮依赖的途径，即直接使血管平滑肌细胞舒张。原发性高血压病患者和自发性高血压大鼠血浆中 CGRP 水平明显降低，给予外源性 CGRP 能有效地降低血压。此外，CGRP 能够抑制血管平滑肌细胞增殖，可用于血管再狭窄的治疗。

3. 肾上腺髓质素　肾上腺髓质素（adrenomedullin，AdM）具有扩张血管、降低血压，抑制血管平滑肌细胞增殖，利钠、利尿和抑制肾素-血管紧张素-醛固酮系统的作用。在各种原因的高血压病患者，其血浆中肾上腺髓质素水平均明显升高，而且升高幅度与高血压的严重程度呈正相关。目前认为，高血压病患者肾上腺髓质素水平升高与交感神经张力升高、肾素-血管紧张素-醛固酮系统过度激活以及体液容量增加等因素有关。静脉给予肾上腺髓质素或转染该基因表达载体可有效降低血压，而且降压迅速，作用持久。肾上腺髓质素有可能成为治疗高血压、心功能衰竭以及心肌梗死等心血管疾病的新药物。

三、一些气体分子是重要的心血管活性物质

在心血管系统中，作为内源性生物活性物质的小分子气体信号分子（gaseous signal molecule）有一氧化氮（nitric oxide，NO）、一氧化碳（carbon monoxide，CO）和硫化氢（hydrogen sulfide，H_2S），它们分别由一氧化氮合酶（nitric oxide synthase，NOS）、血红素加氧酶（heme oxygenase，HO）和胱硫醚-r-裂解酶（cystathionine lyase，CSE）催化生成，并与相应的酶构成 NO/NOS 体系、CO/HO 体系和 H_2S/CSE 体系。在血管平滑肌细胞，NO 和 CO 的作用靶点为可溶性鸟苷酸环化酶和大电导型钙依赖性钾通道，H_2S 直接开放 ATP 敏感性钾通道（K_{ATP}通道）。三种气体分子对血管平滑肌细胞的效应表现为舒张血管平滑肌，扩张血管，调节血压，抑制血管平滑肌细胞增殖和血管重塑。

NO/NOS 体系失调是高血压发病的重要原因之一，应用 NOS 抑制剂阻断 NO 合成已作为制备高血压动物模型的方法之一。在不同类型的高血压中，CO/HO 体系的变化不同。在自发性高血压大鼠，CO/HO 体系在高血压前期即开始下调，在高血压进展期血压升高时，CO/HO 体系显著下调，至高血压稳定期，CO 才代偿性上调。给自发性高血压大鼠补充外源性 CO 可使动物血压下降，抑制正常大鼠内源性 CO 合成后，动物血压升高，这些均说明 CO/HO 体系下调是高血压发生发展的因素之一。在自发性高血压大鼠和用 NOS 抑制剂诱发的高血压动物模型中，H_2S/CSE 体系也明显下调，用 CSE 抑制剂阻断内源性 H_2S 生成后，大鼠血压上升，补充外源性 H_2S 后血压下降，提示 H_2S/CSE 体系异常也是高血压发病机制之一。

第二节　动脉粥样硬化的发病机制与分子基础

动脉粥样硬化（atherosclerosis，AS）是大、中等动脉的一种疾病，主要表现为大、中等动脉的一些部位内膜增厚。内膜下脂质沉积，单核及淋巴细胞浸润；中膜平滑肌细胞迁移至内膜并大量增殖；单核细胞和平滑肌细胞摄取脂质变成泡沫细胞，进而形成动脉粥样斑块。粥样斑块突出于动脉管腔造成血管狭窄，若斑块破裂诱发血栓形成，则会危及生命。动脉粥样硬化的发生是一个非常复杂的过程，多种机制参与发病，如脂质浸润、炎症

反应、损伤应答、氧化应激、载脂蛋白及其受体基因突变等。动脉粥样硬化虽然是多种因素共同作用的结果，但血管平滑肌细胞异常增殖和迁移在动脉粥样硬化斑块形成过程中具有重要作用。

一、血管平滑肌细胞转化和增殖的细胞外信号

血管平滑肌细胞是非终末分化细胞，其表型具有可塑性。血管中膜处于分化状态（收缩表型）的平滑肌细胞在致动脉粥样硬化因素的作用下，可返回去分化状态（合成表型），即发生表型转化（phenotype modulation）（图 14-1），并再进行增殖。血管平滑肌细胞表型转化和增殖受多种因素的影响。

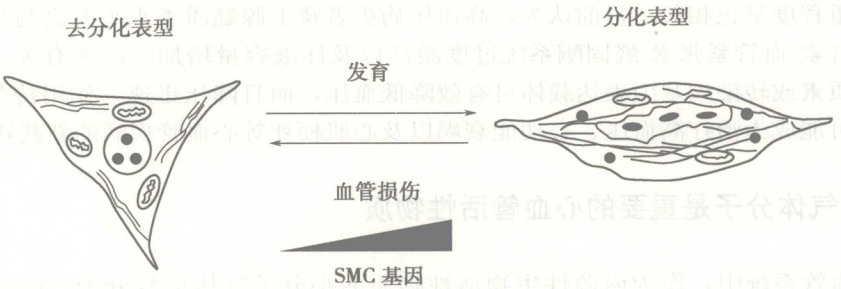


图 14-1　血管平滑肌细胞表型重塑模式图

血管内皮受到致动脉粥样硬化因素的损伤时，循环中的单核细胞和血小板黏附于受损伤的内膜下，在释放大量细胞因子和生长因子的同时，还刺激血管平滑肌细胞分泌生长因子和细胞因子。这些因子通过与血管平滑肌细胞上的相应受体结合，触发平滑肌细胞表型转化和增殖相关的信号转导，诱导即刻早期基因，如 *c-FOS*、*c-JUN* 和 *c-MYC* 表达，进而促使平滑肌细胞表型转化，并从中膜向内膜下迁移及大量增殖，构成动脉粥样硬化斑块组织中的主要细胞成分。与血管平滑肌细胞表型转化和增殖有关的物质总结于表 14-2。

表 14-2　与血管平滑肌细胞表型转化和增殖有关的物质

生长因子和细胞因子
EGF、bFGF、PDGF、VEGF、NGF、TGF-α、TGF-β、IGF、IFN-γ、IL-1β、IL-2、IL-6、TNF-α
血管活性物质
Ang Ⅱ、钠尿肽、ET、AdM、NO、CO、H_2S、肾上腺素/去甲肾上腺素、LTB4、PGI2、PGE1、5-HT、P 物质、加压素
细胞外基质
纤连蛋白、玻连蛋白、层粘连蛋白、胶原、腱生蛋白、骨桥蛋白、血小板反应蛋白

基质金属蛋白酶（matrix metalloprotease，MMP）是分解细胞外基质（extracellular matrix，ECM）的蛋白酶类的总称，其活性依赖于 Zn^{2+} 和 Ca^{2+} 离子。目前已发现的 MMP 有 20 多种。血管平滑肌细胞与 ECM 的相互作用不但决定它们对生长因子和细胞因子的反应能力，而且还影响了平滑肌细胞在收缩表型与合成表型之间的相互转换。ECM 被 MMP 降解对启动平滑肌细胞迁移与增殖是必要的。平滑肌细胞增殖和 ECM 合成与降解之间的失衡，是导致斑块形成的重要原因。临床研究证实，不稳定心绞痛病人动脉粥样硬化斑块组织中的 MMP 水平较稳定型心绞痛患者明显升高，说明斑块局部 MMP 表达增高可能是

斑块不稳定的分子生物学基础。

二、控制血管平滑肌细胞表型的转录因子

血管平滑肌细胞表型受基因表达程序所控制，而基因表达程序取决于特定转录因子的激活。目前已知，调节血管平滑肌细胞表型的转录因子有血清效应因子、心肌素、同源异型蛋白等。

1. 血清效应因子　血清效应因子（serum response factor，SRF）所识别的顺式元件称为 SRE，其中心部位含有 $CC(A/T)_6GG$ 序列，亦称为 CArG 框。在骨骼肌、心肌和平滑肌细胞中特异性表达的基因，其启动子区域大多含有 CArG 框，特别是在分化型血管平滑肌细胞中所表达的基因，在转录起始点上游 12 kbp 内均含有 1 个或几个 CArG 框。SRF 与 CArG 框相互作用是血管平滑肌细胞基因特异性表达所必需的。血管平滑肌细胞表型转化及异常增殖与平滑肌细胞特异性基因表达下调有关。血管平滑肌、心肌和骨骼肌细胞基因特异性表达的条件之一是 SRF 高表达。SRF 在 3 种肌细胞中均有的现象表明，该转录因子只是决定肌细胞基因的表达活性，而不同类型肌细胞基因表达的特异性则是由 SRF 辅激活因子（coactivator）来决定的。

2. 心肌素　心肌素（myocardin）因其在心肌特异性表达及其为心肌基因表达所必需而得名。心肌素属于含 SAP（SAF-A/B、acinus、PIAS）结构域的核蛋白家族成员。除含有 SAP 结构域外，心肌素分子中还含有碱性域、亮氨酸拉链、转录激活结构域（transcription activation domain，TAD）。碱性域是心肌素与 SRF 相互作用的结构域，TAD 发挥转录激活作用，这两个结构域的缺失可使心肌素丧失生肌活性。心肌素分子中的亮氨酸拉链结构介导其同源二聚体化，首先 SRF 在两个相邻的 CArG 框上形成同源二聚体，然后两个心肌素分子结合于 SRF 上形成同源二聚体（图 14-2），这可能是血管平滑肌细胞基因表达依赖于两个 CArG 框的原因。

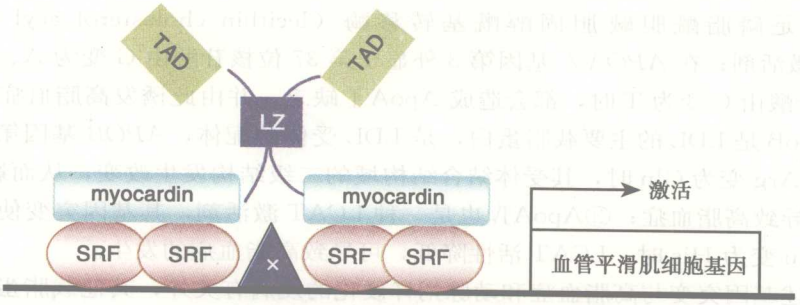


图 14-2　心肌素调节平滑肌基因表达的模式图

心肌素是能启动平滑肌分化程序的两个转录因子之一。另一个能激活平滑肌分化程序的转录因子是心肌素相关转录因子 A（myocardin-related transcription factor A，MRTF-A），该因子与心肌素之间结构高度相似。

3. 同源异型蛋白　同源异型蛋白（homeoprotein）绝大多数是转录因子，该类蛋白含有高度保守的、由 60 个氨基酸残基构成的螺旋-转角-螺旋结构域，即同源异型结构域。同源异型结构域通过与特定基因的启动子序列结合而实现对靶基因的转录调控作用。

在胚胎期血管平滑肌细胞中，特异性表达的同源异型蛋白有 Prx2、Hox7、Hox9 和 Hex 等；在成年血管壁平滑肌细胞中，特异性表达的同源异型蛋白包括 Prx1（Mhox）、

Gax、Nkx3.2、Hox5、HoxA11 和 HoxB1 等。这些同源异型蛋白通过以多种方式与 SRF 及其他转录因子相互作用而调节平滑肌细胞特异性基因的表达及细胞表型。同源异型蛋白调节血管平滑肌细胞表型的分子机制如下：一是通过直接或间接与 SRF 结合，改变以 SRF 为核心的转录复合物与特定基因表达调控序列结合的特异性，或是提高或降低转录复合物对特定基因调控序列的亲和力，进而启动平滑肌细胞分化相关基因的表达；二是通过提高平滑肌细胞分化相关基因启动子周围的组蛋白乙酰化水平，促进染色质重塑，进而提高转录复合物与相应序列的亲和力，从而起到对血管平滑肌细胞基因的转录调控作用。

三、高脂血症是动脉粥样硬化发生发展的重要原因

由于遗传和环境因素的影响，血浆脂蛋白的结构和代谢发生异常，血液中某种或几种脂质成分升高时称为高脂血症（hyperlipidemia）。高脂血症与动脉粥样硬化的发病有着密切关系。

在 4 种脂蛋白中，低密度脂蛋白（low density lipoprotein，LDL）在动脉粥样硬化发生发展中起非常重要的作用。在正常情况下，血浆中的 LDL 与细胞膜表面的 LDL 受体结合后被吞入细胞，在溶酶体酶的作用下，蛋白质降解为氨基酸，胆固醇酯水解为游离胆固醇和脂肪酸。当机械、促炎因子等因素造成内皮损伤时，可促使 LDL 进入内皮下层组织，处于内皮下的 LDL 易于被氧化修饰。氧化修饰的 LDL 可激活内皮细胞和循环中的单核细胞，使单核细胞黏附于血管内皮，并转变为巨噬细胞，后者摄取脂质而成为泡沫细胞。

LDL 受体基因含有 18 个外显子和 17 个内含子。目前已发现 70 多种 LDL 受体基因的突变。这些突变包括点突变、小片段或大片段插入、缺失或重排。LDL 受体基因突变可引起不同类型的受体缺陷，致使血液中的 LDL 胆固醇（LDL-C）不能正常代谢，血浆 LDL-C 水平显著升高。

载脂蛋白（apolipoprotein，Apo）及其受体基因的缺陷也是高脂血症发生的重要原因。较重要的有：①ApoAⅠ是高密度脂蛋白（high density lipoprotein，HDL）的主要载脂蛋白，也是磷脂酰胆碱-胆固醇酰基转移酶（lecithin cholesterol acyl transferase，LCAT）的激活剂；在 *APOA*Ⅰ基因第 3 外显子第 37 位核苷酸由 G 变为 A、第 4 外显子第 84 位核苷酸由 C 变为 T 时，都会造成 ApoAⅠ缺乏，并由此诱发高脂血症和动脉粥样硬化；②ApoB 是 LDL 的主要载脂蛋白，是 LDL 受体的配体，*APOB* 基因第 3 500 位密码子编码的 Arg 变为 Gln 时，其受体结合结构域的二级结构发生改变，从而影响 LDL 与受体结合，导致高脂血症；③ApoAⅣ也是一种 LCAT 激活剂，其基因突变使第 360 位的氨基酸由 Glu 变为 His 时，LCAT 活性降低，可导致高脂血症的发生。

除了上述基因突变与高脂血症和动脉粥样硬化的发病有关外，其他载脂蛋白以及脂蛋白脂酶（lipoprotein lipase，LPL）基因多态性也与高脂血症和动脉粥样硬化的发病有关。

四、慢性炎症反应贯穿于动脉粥样硬化发生发展的全过程

动脉粥样硬化发生发展的不同阶段均存在慢性炎症反应。在早期，脂质在动脉壁上的沉积，以及动脉分支处受血液湍流的冲击，使血管内皮受损。受损的血管内皮细胞表达各种黏附分子，后者通过与其配体结合，介导内皮细胞与单核细胞之间的黏附。在趋化因子和单核/巨噬细胞集落刺激因子（macrophage colony stimulating factor，M-CSF）的作用下，与内皮细胞黏附的单核细胞向内膜下迁移，活化成为巨噬细胞，并释放 IFN-γ、TNF-α 和 IL-1β 等炎性因子，继而刺激血管平滑肌和内皮细胞的炎性反应，*COX-2* 和

NOS 基因表达增强，前列腺素和 NO 合成增加，以及血清中 C 反应蛋白（C-reactive protein，CRP）水平升高等。

炎性细胞一旦在动脉壁中聚集、活化后，就会参与斑块局部的免疫应答反应，进一步促进动脉粥样硬化斑块的形成和发展。斑块中的炎性细胞除了释放细胞因子和生长因子外，还和动脉壁中的平滑肌细胞合成的细胞外基质一起在斑块的外层包上一层纤维帽，从而形成纤维斑块。随着病变的发展，最后纤维帽内部的细胞发生坏死，与脂质混合形成粥糜样黄色物质，即典型的粥样斑块。

在动脉粥样硬化发生早期，以及斑块形成、发展和合并症发生过程中，均伴有 NF-κB 的表达上调和活化。NF-κB 是调节炎性基因表达的转录因子，可启动黏附分子、趋化因子、细胞因子、生长因子等多种炎症反应相关基因的表达。这进一步说明，炎症反应贯穿于动脉粥样硬化发生和发展的全过程。

五、活性氧促进动脉粥样硬化的发生与发展

当机体抗氧化能力不足以防止 ROS 引起氧化性损伤时，即发生氧化应激（oxidative stress）。ROS 主要通过对脂蛋白的氧化作用和诱导炎性介质基因表达来促进动脉粥样硬化的发生与发展，在高血压、糖尿病、高脂血症等引起动脉粥样硬化的发病过程中起重要作用（图 14-3）。

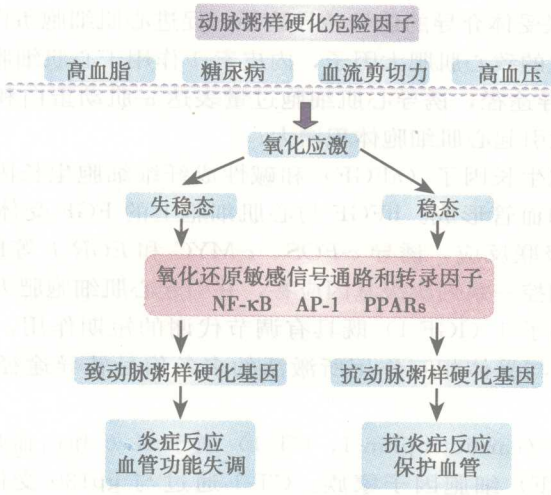


图 14-3　动脉粥样硬化危险因子与氧化还原敏感基因之间的相互联系

ROS 对炎性介质基因表达的影响是通过调节转录因子活性来实现的。细胞内的 ROS 既可直接氧化修饰转录因子，又可通过磷酸化/去磷酸化修饰的级联反应调节转录因子活性。激活蛋白-1（activator protein-1，AP-1）和 NF-κB 是受氧化还原状态影响的主要转录因子，在血管细胞中，ROS 既可通过使 AP-1 磷酸化增强其 DNA 结合活性，还可通过使 NF-κB 抑制蛋白（inhibitor-κ binding，IκB）降解导致 NF-κB 活化。近年发现，过氧化物酶体增殖剂激活受体（peroxisome proliferator-activated receptor，PPAR）也是氧化还原敏感性转录因子，并参与动脉粥样硬化的炎症反应过程。

沉积在动脉内膜下的 LDL 失去血液中抗氧化物质的保护，其分子中的不饱和脂肪酸易于被脂氧化酶和髓过氧化物酶所氧化。氧化型 LDL 具有更强的致动脉粥样硬化作用。

第三节　心肌肥厚的分子机制

心肌细胞是一种高度分化的终末细胞，这类细胞不能再进行增殖，只能有细胞体积的增大。因此，从某种意义上来讲，心肌肥厚指的是心肌细胞的肥大。心肌肥厚（cardiac hypertrophy）是高血压、动脉粥样硬化等多种心血管疾病的常见并发症，具有心肌细胞体积增大、心肌间质细胞增殖和结缔组织增生、肌原纤维新形成的肌节增加和胚胎基因再表达等生物学特征。

一、心血管活性肽和生长因子引起心肌肥厚

致心肌肥厚因素，如高血压引起的心脏负荷增加、神经刺激、缺血损伤等均可以促进一些生长因子和心血管活性肽的合成与分泌。这些生长因子和活性肽通过不同的细胞信号转导途径，首先诱导即刻早期基因（如 *c-JUN*、*c-FOS*、*EGR-1*）表达，其后是胚胎基因 *ANF* 和胚胎型肌球蛋白重链（*βMHC*）重新表达，最后是心肌肥大相关基因，如肌动蛋白、肌球蛋白轻链、原肌球蛋白和胶原蛋白等表达增加。这些蛋白的高表达致使心肌细胞体积增大。不同刺激因素诱导基因表达的模式不尽相同，这主要取决于它们启动的信号转导途径。致心肌肥厚的刺激因子主要有以下几种：

（1）在心肌肥厚发生早期，血管紧张素Ⅱ通过与心肌细胞上的Ⅰ型受体（AT1R）相互作用，激活 G 蛋白偶联受体介导的信号转导途径，促进心肌细胞蛋白质的合成。

（2）内皮素-1 是强效的致心肌肥大因子，内皮素-1 作用于心肌细胞上的受体后，激活 PKC 和 Ca^{2+} 两个信号转导途径，诱导心肌细胞过量表达 α-肌动蛋白和胶原蛋白，重新启动胚胎基因表达，并由此引起心肌细胞体积增大。

（3）酸性成纤维细胞生长因子（aFGF）和碱性成纤维细胞生长因子（bFGF），两者均可促进心肌细胞生长和血管形成。bFGF 与心肌细胞上的 FGF 受体（FGFR）结合后，通过酪氨酸激酶磷酸化级联反应，诱导 *c-FOS*、*c-MYC* 和 *EGR-1* 等即刻早期基因表达，这些表达产物可进一步调控一系列次级基因应答，并引起心肌细胞肥大。

（4）胰岛素样生长因子-1（IGF-1）既具有调节代谢的短期作用，又有促进各种细胞增殖的长期作用。IGF-1 与受体相互作用所激活的多条信号转导途径均参与心肌肥厚的发生。

（5）心脏营养蛋白-1（cardiotrophin-1，CT-1）属于 IL-6 和白血病抑制因子（leukemia inhibitory factor，LIF）细胞因子家族。CT-1 通过与 gp130 受体结合，激活 JAK-STAT 途径，促进未成熟心肌细胞的存活和增殖，并诱导心肌细胞肥大，在多种心血管疾病所致的心肌肥厚和心功能衰竭的发病中发挥重要作用。

二、心血管活性肽和生长因子致心肌肥厚效应的信号转导途径

致心肌肥大因子与受体相互作用所触发的信号转导途径有 MAPK 途径、JAK/STAT 途径、Ca^{2+} 介导的信号途径和 PI3K 途径。

（1）MAPK 途径：MAPK 家族的 ERKs、JNK 和 p38MAPK 三个成员都参与心肌肥厚的信号转导。MAPK 信号途径经多个激酶的级联反应将信号传递到胞核，通过使转录因子磷酸化激活基因表达。在高血压导致心肌肥厚的发生过程中，ERK1/2 在心肌肥厚的形成阶段起关键作用，JNK 和 p38MAPK 在心肌肥厚代偿期活性增加，在失代偿期则活

性降低，以致失活。

（2）JAK/STAT 途径：心营养蛋白和 Ang Ⅱ与相应的受体结合后导致 JAKs 激活，催化 STAT 的磷酸化，使其发生二聚体化后转移进入细胞核，结合到靶基因启动子的顺式元件上并激活基因转录。研究表明，Ang Ⅱ一方面通过 AT1R 激活 JAK2、STAT3/STAT6 和 STAT5/STAT6，另外还能诱导 Ang Ⅱ原基因表达，引起自分泌性 Ang Ⅱ的大量产生。大量的 Ang Ⅱ激活 JAK2-STAT1 途径可导致心肌细胞凋亡。

（3）Ca^{2+}介导的信号途径：Ang Ⅱ与受体结合后，通过 PLC 活化生成 IP3，IP3 促进心肌细胞内对 IP3 敏感的 Ca^{2+}贮存库释放 Ca^{2+}至胞浆，使胞浆内 Ca^{2+}浓度升高。DG 激活 PKC。胞浆内 Ca^{2+}升高和 PKC 激活可激活即刻早期基因转录因子，如 *c-FOS*、*c-MYC*、*c-JUN*、*EGR-1* 等，进一步诱导心肌细胞表达胚胎型基因。Ca^{2+}对心肌细胞肥大的调节还涉及 Ca^{2+}/钙调蛋白依赖性蛋白激酶的激活。

第四节　心律失常的分子机制

心律失常（arrhythmia）是指心脏节律发生紊乱，并引起血流动力学的显著异常。心律失常既可以原发于先天性冲动形成和传导异常，也可以继发于各种心脏及全身性疾病。心律失常大多是环境因素和遗传因素共同作用的结果。引起心律失常的间接原因有高血压、冠心病、心肌病、先天性心脏病和瓣膜病等，这些疾病均和遗传因素有关；心肌细胞离子通道基因突变是引起心律失常的直接原因。近年，在遗传性或家族性心律失常的分子机制研究方面取得较大进展，本节重点介绍长 QT 综合征（long QT syndrome，LQTS）和心房颤动（以下简称为房颤）的分子生物学机制。

一、编码心肌离子通道蛋白的基因突变引起 LQTS

LQTS 是在心电图上表现为 QT 间期延长，易发生尖端扭转性室速和室颤等室性心律失常，在临床上表现为晕厥和猝死的一组综合征。

（一）LQTS 的分型

LQTS 有两种遗传方式，一种是常染色体显性遗传，称为 Romano-Ward 综合征，只存在心肌疾患的表现型；另一种为常染色体隐性遗传，称为 Jervell-Lange-Nielsen 综合征，该类型的心肌病变与先天性耳聋并存。

表 14-3　LQTS 类型、致病基因、染色体定位及发生突变的通道蛋白

LQTS 亚型	变异基因	基因位点	编码产物	发病机制
LQT1	*KCNQ1*（*KvLQT1*）	11p15.5	缓慢延迟整流钾通道 α 亚单位	心肌复极时间延长
LQT2	*HERG*	7q35-36	快速延迟整流钾通道 α 亚单位	心肌复极时间延长
LQT3	*SCN5A*	3p21-24	电压门控钠通道 α 亚单位	钠离子电流持续时间延长
LQT4	*ANK2*	4q25-27	锚蛋白 B	Ca^{2+}通道活性降低
LQT5	*KCNE1*（mink）	21q22.1-22.2	缓慢延迟整流钾通道 β 亚单位	改变激活通道的电压依赖性，加速通道失活
LQT6	*KCNE2*（MiRP1）	21q22.1-22.2	快速延迟整流钾通道	影响心肌细胞复极过程
LQT7	*KCNJ2*	17q23.1-24.2	内向整流钾通道	心肌动作电位末期时相延长

（二）LQTS 的分子生物学机制

LQTS 是由 6 个编码离子通道的基因突变和 1 个影响钠通道动力学的细胞骨架基因突变引起的。

1. *KCNQ1* *KCNQ1* 基因编码 676 个氨基酸组成的电压门控钾通道（Kv）的 α 亚单位，是最早发现的与 LQTS 相关的基因，因此，又将其命名为 Kv LQT1。KCNQ1 的正常功能需要由 β 亚单位 KCNE1 辅助才能实现。KCNQ1 和 KCNE1 相互作用，装配成缓慢激活钾通道。*KCNQ1* 基因突变（如 R190Q、G269D、Y315S、R555C 等）后，与 KCNE 形成的钾通道则功能丧失，造成心肌复极时间延长和 LQTS。

2. *HERG* *HERG*（human ether a-go-go-related gene）基因编码 1 159 个氨基酸组成的快速激活延迟整流钾通道（Ikr）的 α 亚单位。目前发现，至少 *HERG* 基因的 80 个点突变与 LQTS 的发生有关。这些突变多数造成单个氨基酸置换，少数由多个变异共同作用引起 LQTS 的发生。

3. *SCN5A* *SCN5A* 基因编码 2 016 个氨基酸组成的电压门控钠通道的 α 亚单位。该通道激活与心肌细胞快速除极有关。*SCN5A* 基因突变可以引起多种心律失常，如 LQT3、Brugada 综合征、进行性和非进行性心脏传导阻滞和特发性心室颤动等。到目前为止，已发现的 *SCN5A* 基因突变位点有 105 个，其中 30 多个位点引起 LQT3。*SCN5A* 基因突变使钠通道失活延迟，复极和动作电位的时程延长。由于动作电位时程的异常延长，使早期后除极及触发活动增加，并诱发尖端扭转型室速。

4. *KCNE1* *KCNE1* 基因编码 130 个氨基酸组成的 mink 蛋白，是 LQT5 的致病基因。mink 蛋白是 Iks 的 β 亚单位，通过其 C 端和 KCNQ1 装配成缓慢激活延迟整流钾通道，形成 Iks 电流。在 *KCNE1* 基因发现的突变较少，目前已有 S74L 和 D76N 两种错义突变。这两种突变通过改变 Iks 的电压依赖性和加速通道失活，导致 Iks 的电流减少，复极延迟，使 QT 间期延长。

5. *KCNE2* *KCNE2* 基因编码由 123 个氨基酸组成的 mink 相关蛋白（MiRP1）。MiRP1 与 HERG 结合，形成完整的快速激活延迟整流钾通道（Ikr）。该基因突变通过引起心肌细胞动作电位钾离子流的改变而影响心肌细胞的复极过程，是 LQT6 的致病基因。迄今，已发现 Q9E、M54T、I57T 和 A116V 与 LQTS 发生有关。

6. *KCNJ2* *KCNJ2* 基因编码由 5397 个氨基酸组成的对 ATP 敏感的内向整流钾通道，该基因突变导致一种罕见的遗传性 Andersen 综合征，表现为周期性瘫痪、以 QT 间期延长为特征的室性心律失常（LQT7），以及骨骼肌发育障碍。

二、多种因素参与房颤的发生

心房颤动（atrial fibrillation，AF）指心房快速的无序激动，心房失去收缩功能，是临床上最常见的持续性快速心律失常，简称房颤。房颤多与器质性心脏病，如心功能衰竭、心肌病、冠心病、高血压等相伴发，但也可没有心脏病基础，后者称作孤立性房颤或特发性房颤，这种房颤具有一定的家族遗传倾向。

房颤发病的分子生物学机制较为复杂，其候选基因主要有心房电活动相关基因、自主神经功能调节相关基因、心房发育与结构重构相关基因等。其中，缓慢延迟整流钾通道亚单位、β-肌球蛋白重链、肌钙蛋白 T 基因突变与家族性房颤的发病有关；线粒体 DNA 碱基的缺失和突变与散发性房颤的发病有关。

1. *KCNQ1* *KCNQ1* 基因除与 LQTS 有关外，还是一个已知的房颤致病基因。*KCNQ1* 基因编码区第 418 位的 A 被 G 取代时，则导致第 140 位的丝氨酸变为甘氨酸

（S140G），家族性房颤患者均存在 S140G 错义突变。但是，单纯 S140G 突变并不能使钾电流增加，只有当其与 β 亚单位 *KCNE1* 共表达形成钾离子通道时，才能增加电导性，缩短心房肌动作电位时间和维持房颤。另外，*KCNQ1* 也可与 *KCNE2*、*KCNE3* 形成功能性钾离子通道。S140G 突变的 *KCNQ1* 与 *KCNE2* 共表达形成的钾离子通道使去极化心肌细胞钾外流增加，不应期缩短；使超极化心肌细胞钾内流增加，稳定静息电位。而与 LQTS 有关的突变（R190Q、G269D、Y315S、R555C）与 *KCNE2* 共表达形成的钾离子通道则导致通道功能完全丧失，可见，S140G 突变是引起钾电流增加的主要因素。

2. 连接蛋白　相邻心肌细胞间的连接结构称为缝隙连接，缝隙连接由连接蛋白构成。在哺乳动物心脏存在多种连接蛋白，如 Cx37、Cx40、Cx43、Cx45、Cx46 和 Cx50。连接蛋白异常可引起传导速度的异质性，诱发房颤的发生。已经证明，Cx40 与术后房颤的发作有关。

正如 LQTS 一样，房颤的分子基础已经超出离子通道，炎症、退行性变、氧化应激、线粒体 DNA 突变、各种信号转导系统等也可能与房颤发病有关。要探明其发病机制可能还需跨越漫长的道路。

（温进坤　郑　斌）

第十五章　内分泌及代谢疾病的分子机制

关键知识点

- 激素作为高等生物细胞间重要的可溶性化学信号之一，是协调整合细胞间功能以适应外界变化的重要通讯方式。激素在高等动物体内的作用方式都具有网络调节特点，保证机体对环境的适应具有稳态平衡的能力。
- 内分泌细胞群的发育需要诱导细胞分化的信号和细胞对诱导信号的正确反应。诱导细胞分化的刺激信号存在于胚胎环境中的邻近细胞。在内分泌系统器官形成中最为重要的转录因子是含有发育同源结构域的蛋白分子和核受体蛋白这两类转录因子。
- 在成熟个体，激素在基因表达氨基末端均有信号肽结构、需经过多层次翻译后修饰、可以在细胞内精确的定位和分泌。内分泌系统所有疾病都或归因于激素自身的结构和含量异常，或由于靶细胞的敏感性异常。

关键词： 信号肽；发育同源选择因子；2 型糖尿病相关基因

内分泌系统由分泌各种微量化学物质——激素的不同细胞群（内分泌腺）构成，从中胚层和内胚层发育而来。激素作为高等生物细胞间重要的可溶性化学信号之一，是协调整合细胞间功能以适应外界变化的重要通讯方式。激素通过血液循环到达靶细胞，与相应的受体相结合，经由复杂的信号转导过程，最终通过影响靶细胞的代谢、分裂、分化等行为而发挥广泛的全身性作用。内分泌系统与神经系统、免疫系统相配合，维持机体内环境的平衡。各系统间有一套完整的互相制约、互相影响的复杂的正负反馈系统，保持机体内环境在外环境条件变化时的稳定，这是维持生命和保持种族延续的必要条件。

任何一种内分泌细胞的功能失常所致的一种激素分泌过多或缺乏，均可引起相应的病理生理变化。本章讨论内分泌系统的分子生物学特点和内分泌系统疾病的分子机制。

第一节　内分泌代谢系统的分子生物学特点

一、内分泌系统的进化

高等生物体内的大分子、细胞器、细胞、组织和器官在空间上是相互隔离的，生物体与环境之间更是如此，所以生物体在新陈代谢时，不但有物质与能量的变化，即存在物质流与能量流之外，还要有信息的流动来控制物质流和能量流的有序进行。神经系统、内分泌系统和免疫系统都是信息流的重要收发系统。

激素信号在生命进化早期就已经出现。在细菌、酵母等低等生物中已经可以找到肽类激素的进化痕迹。例如，细菌中的 α 因子作为一种外激素，在分子结构上与哺乳动物的促黄体激素释放激素（luteinizing hormone-releasing hormone，LHRH）类似；血管活性肠肽（vasoactive intestine peptide，VIP）与一种蠕虫的胰腺样肽结构类似；胰岛素（insulin）和 ACTH（adrenocorticotropin）样分子在四膜虫中已经存在。

内分泌系统的出现是生物适应环境而不断变异的结果，实质上是细胞通讯系统（信息流系统）从低级到高级的发展完善过程。单细胞生物与外环境直接交换信息，外环境的变

化将导致细胞对自身功能的调节，以利于更好地生存和繁衍后代。多细胞生物中的每一个细胞不仅需要适应其所在环境的变化，更重要的是细胞与细胞之间需要在功能上的协调统一。按照这种需求，最原始的方式是细胞与细胞的直接联系，表现为物质的直接交换，或者是通过细胞表面分子的相互作用而实现的信息交流，这样的调节方式直至现在仍然是高等动物细胞分化、个体发育及实现协调的整体功能适应的重要方式之一，例如各种免疫细胞之间广泛的相互作用。但是，对于远距离细胞之间的功能协调则必须有可以长距离作用的信号，激素就是适应这种需求而出现的信号系统。

从进化角度看，内分泌激素应该是在局部发挥作用的旁分泌信号（如细胞因子）的基础上发展而来的。细胞因子调节局部细胞与细胞之间的相互作用，而激素则负责个体内器官间的调整。

原始的内分泌系统是以神经分泌形式为主，即由神经元分泌激素，或者直接分泌到循环中，或者储存在神经与血管的接触结构中，在需要时分泌。大部分有发育良好的神经系统和循环系统的动物都有内分泌系统存在。甲壳类、节肢类和脊椎动物的内分泌系统十分类似，是趋同进化的一个典型例证。脊椎动物的内分泌系统由内分泌腺体和散在于上皮组织的内分泌细胞组成。

内分泌腺体在胚胎发育中来自于所有的胚胎组织层面（内胚层、中胚层和外胚层）。内分泌腺分泌产物的类型取决于其所来源胚胎组织。外胚层和内胚层来源的腺体产生的激素主要是肽类和氨基酸及其衍生物；而中胚层来源的腺体产生的激素主要是脂类。

由于环境的复杂性，机体对环境的适应必须具有稳态平衡的能力。这种稳态平衡是通过网络性调控而实现的。无论是细胞因子还是激素，在高等动物体内的作用方式都具有网络调节特点。这种网络式调节体现在：一种细胞因子或激素的作用始终会受到其他细胞因子或激素的影响，或抑制、或促进；发出信号的细胞随时又受到其他细胞发出的信号的调节。这种网络的存在使得机体内的细胞因子或激素的作用都具有一定程度的冗余和代偿性，单一缺陷不会导致对机体的严重损害。激素信号在细胞内的转导过程已在本书第八章详细论述。信号在细胞内的转导在另一个层次上形成复杂的网络系统。

生物的进化始终是围绕着能量的有效利用和储存这一核心进行的，内分泌系统作为体内一个重要的调节系统，同样在代谢调节方面具有中心作用。内分泌系统的疾病也往往主要体现在能量代谢调节的异常。

二、内分泌系统发育的分子调控

生命体的生长发育主要受遗传信息及环境变化信息两方面的控制。体内所有的细胞都携带着同样的遗传信息。遗传基因决定了个体发育的基本模式，机体外环境和内环境两个方面的信息则给予个体遗传基本模式的完整表现以足够的刺激信号。

内分泌细胞群的发育同样需要诱导细胞分化的信号和细胞对诱导信号的正确反应。诱导细胞分化的刺激信号存在于胚胎环境中的邻近细胞，如果将胚胎细胞移动到另一位置，将改变细胞的分化命运。诱导细胞分化的信号包括存在于细胞表面的一些分子如 Notch 等，细胞因子如 BMP、EGF、FGF、TGF 等。细胞的反应则主要是经由转录因子而实现。

哺乳动物的器官形成是从一个胚原细胞形成具有特定表型细胞的程序化过程。目前认为，这种程序实质是由转录因子的时空表达程序所决定的。器官形成是由发育过程中某一时空阶段，一套转录因子的特定表达所决定的。所以，基因表达谱的差异是细胞分化和器官发育的根本原因，诱导细胞分化的所有信号都将在靶细胞内最终作用于组织特异性转录因子。在内分泌系统器官形成中，转录因子同样扮演着关键角色，其中最为重要的是含有

同源异形结构域（homeodomain）的蛋白分子和核受体（脂溶性化学信号受体）蛋白这两大类转录因子。

含有发育同源结构域的蛋白分子亦称为同源异形选择因子（homeotic selector），是一个从果蝇到人类都十分保守的基因家族。这些基因及受其激活的下游靶基因的时空有序表达，决定了组织、肢体、器官等的发育顺序。内分泌系统的发育与这一类转录因子的作用密切相关。同时各种类固醇激素、甲状腺素、性腺激素、维甲酸等对内分泌系统的正常发育也是必需的，它们的作用是通过核受体类转录因子实现的。这些关键的转录因子如果发生突变，也可以导致相应器官发育异常。对于这些转录因子在发育中的作用，或者是通过相应的人类或小鼠遗传病的基因定位分析而认识，或者是通过基因靶向剔除后的表型变化而确定的。这里以部分目前已经有所了解的内分泌腺体发育过程为例，说明特异性转录因子的作用。

（一）胰腺发育相关基因

胰腺的发育过程不仅有内分泌细胞与外分泌细胞的分化，还要有不同的内分泌细胞群（分泌胰高血糖素的 A 细胞、分泌胰岛素的 B 细胞、分泌生长抑素的 δ 细胞和分泌胰多肽的 PP 细胞）需要精确的分化。这些分化和发育过程同样是特异性转录因子控制的基因时空差异性表达所致。图 15-1 是胰腺发育阶段及其相关转录因子的示意图。

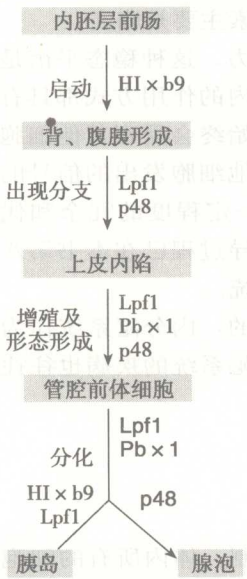


图 15-1　胰腺发育及关键转录因子

与胰腺发育相关的重要的转录因子主要有以下几种：

（1）Hlxb9：Hlxb9（homeobox gene HB9，Hlxb9）在成人胰腺表达。*HLXB9* 在小鼠的同源基因 *hlxb9* 在小鼠胚胎发育及成熟的 B 细胞中均有表达。*hlxb9* 基因剔除小鼠纯合子表现为胚胎期背胰发育障碍，在残余胰腺中胰岛素分泌细胞明显减少。

（2）Ipf1：Ipf1（insulin promoter factor 1，IPF1）亦称为 STF1（somatostatin transcription factor）。最初发现其可以与胰岛素基因启动子及生长抑素启动子特异性结合，其后发现与胰腺发育相关。Ipf1 是胰腺发育过程中控制胰腺形成的主要转录因子。*ipf1* 基因剔除小鼠表现为胰腺发育障碍。

（3）Ptf1-p48：PTF1-p48（pancreas transcription factor 1），亦称为胰岛-十二指肠分化基因（islet duodenum homeobox-1，IDX-1），是胰腺转录因子的 α 亚单位，在胰腺的内分泌和外分泌组织的形成中均具有重要调节作用。缺乏 PTF1-p48 的胰腺上皮细胞将不再向胰腺发育，而是返回到肠上皮命运。先天性无胰腺患儿是 PTF1-p48 基因突变纯合子，其双亲和祖先往往为 PTF1-p48 突变的杂合子，常在成年期成为Ⅱ型糖尿病患者。

（4）Pbx1：Pbx1（pre-B-cell leukemia transcription factor 1，Pbx1）也属于 homeodomain 转录因子，最初在人的急性白血病的一种染色体易位突变中被发现，后来又认识到 Pbx1 在小鼠、果蝇和鸡胚胎肢体发育中的作用。体外研究发现 Pbx1 可以调节 Ibf1 的活性，进而影响胰腺的发育。Pbx1 基因剔除纯合子小鼠在胚胎 15～16 天死亡，胚胎形态观察可见严重胰腺发育不全，杂合子小鼠成熟个体有胰岛功能不全。

（二）肾上腺、性腺发育相关基因

肾上腺髓质来源于胚胎期的神经嵴。神经嵴细胞迁移到后肾的上方后，包绕神经嵴细胞的间叶细胞最后形成胚胎型肾上腺皮质的球状带、束状带和网状带，出生后，再为成熟的肾上腺皮质所替代。

胚胎期肾上腺的生长和发育受到 ACTH 和局部其他生长因子（如 IGF-Ⅱ和 bFGF）

的直接影响。出生后，胚胎型肾上腺皮质细胞在活化素（垂体合成的性激素）和 TGF-β 的作用下进入凋亡程序。

目前认为有两种转录因子与肾上腺、性腺及下丘脑的腹侧中央区发育有关。一个是 SF-1（steroidogenic factor-1），可以结合在基因组中的雌激素结合元件的部分区域，该基因的剔除会引起肾上腺和性腺发育不全。此外，垂体-性腺区发生缺失，下丘脑的腹侧中央区也严重发育不全。

另一个相关的转录因子是 DAX-1（Dosage-sensitive sex reversal，Adrenal hypoplasia congential，X-chromosome），是人类先天性低肾上腺素血症的致病基因，位于 X 染色体，呈现剂量依赖性转换。DAX-1 可以结合到一些特殊的基因调控位点，例如与肾上腺-性腺发育相关的基因。DAX-1 主要作用是抑制基因的表达。DAX1 的突变对男性早期性腺发育没有影响，但是肾上腺皮质的发育只能停留在胎儿阶段。

内分泌系统的发育机制研究尽管在最近有了较多进展，但是距离阐明各个内分泌器官发育的起始信号、在不同阶段的分化转折信号、相关的基因表达谱变化和关键的转录因子都仍然在探索之中。研究内分泌系统的发育机制，对于认识内分泌系统疾病的发病机制将具有重要帮助。

三、肽类激素合成与分泌的分子调控特点

肽类激素基因的结构及其表达过程有以下几个特点：

（1）氨基末端均有信号肽结构：由于肽类激素均为分泌性蛋白，为分泌到细胞外之需要，在肽类激素的 N-端均有一段疏水性氨基酸构成的信号肽（引导肽）。信号肽序列负责识别内质网膜中的特殊信号肽识别位点，引导新生肽进入分泌途径。

（2）需经过多层次翻译后修饰：绝大部分肽类激素是以前体形式合成的，在翻译以后经历诸如糖基化、硫化等共价修饰，以形成完成活性所需要的正确构象，尤其是冗余肽的切除对很多激素都是需要的。图 15-2 给出了一些例子说明蛋白裂解这一修饰方式对于激素成熟的必要性。

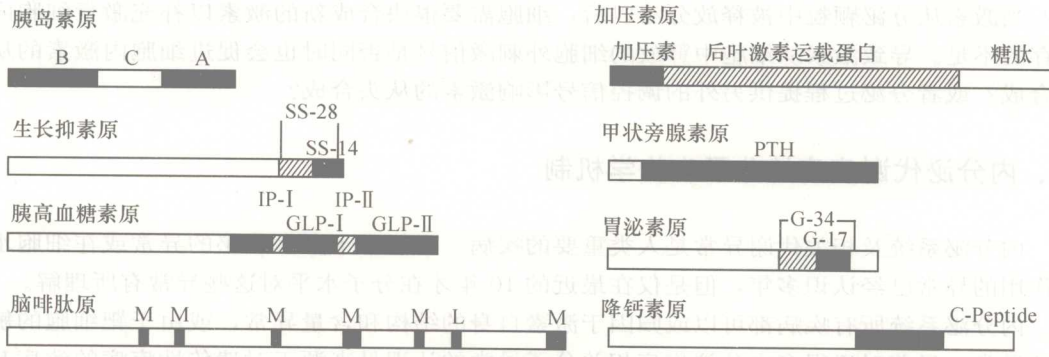


图 15-2　部分激素原结构

（3）需要在细胞内精确的定位和分泌过程：新合成的肽首先进入内质网，形成转运小体，再包装到分泌颗粒内，在需要时释放到细胞外。该过程可见图 15-3。内分泌细胞都具有储存已经合成的激素的作用，在需要时可以很快分泌到细胞外。这一过程始终受到外源信号的调节，与细胞内许多信号转导分子的活化有关，尤其是与那些与细胞骨架重构，细胞内转运小体（transport vesicles）及分泌颗粒运动有关分子的活化密切相关。

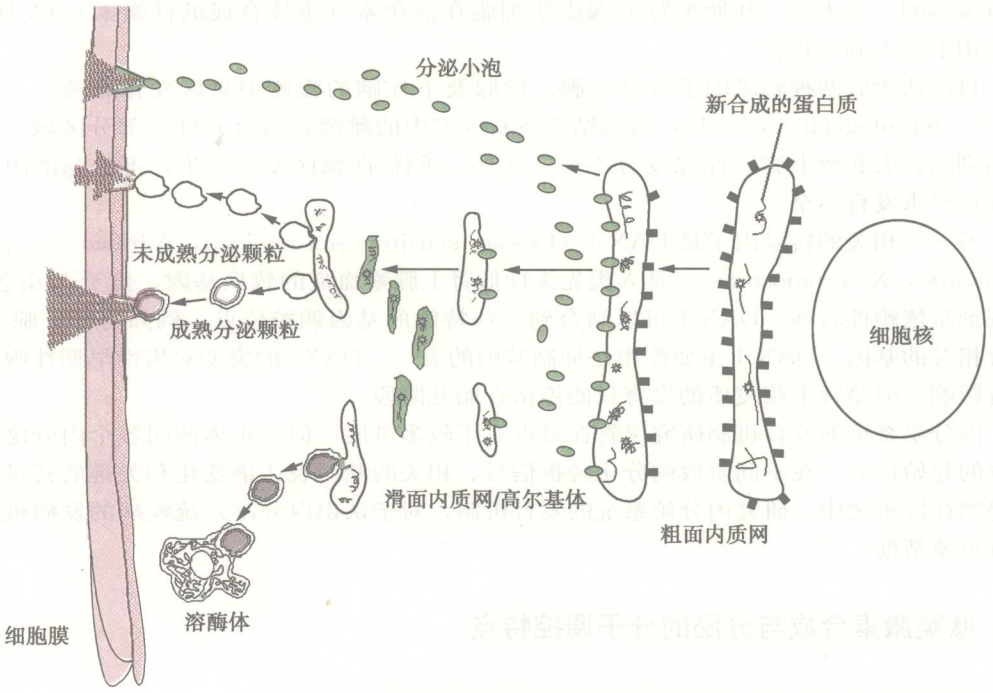


图 15-3　激素合成后分泌机制示意图

（4）需要激素分泌与合成调节的协调：激素的合成与分泌在体内必须有精确的调节，才能完成保持机体稳态平衡的任务。激素合成的调节机制与其他基因表达调控机制是一致的，其基本过程都是细胞所处环境的改变（如细胞之间的相互作用、细胞因子的种类和浓度、细胞营养及代谢物种类和浓度等）作为信号，作用于细胞内的信号转导网络系统，最终改变细胞基因表达调控相关分子（如转录因子、蛋白合成相关分子等）的活性，使得相应激素的合成加快或减慢。

在激素合成与分泌调节中一个尚未解决的问题是合成调节与分泌调节之间如何协调的？当激素从分泌颗粒中被释放分泌以后，细胞需要很快合成新的激素以补充激素细胞内储存的不足。导致激素从细胞中释放的细胞外刺激信号是否同时也会促进细胞内激素的从头合成？或者分泌过程提供另外的调控信号影响激素的从头合成？

四、内分泌代谢疾病的分子生物学机制

内分泌系统及相关代谢异常是人类重要的疾病。人们对于激素分泌的异常或在细胞水平作用的异常已经认识多年，但是仅在最近的 10 年才在分子水平对这些异常有所理解。

内分泌系统所有疾病都可以或归因于激素自身的结构和含量异常，或由于靶细胞的敏感性异常。目前对于很多内分泌疾病相关分子异常的认识是来源于对遗传性疾病的致病基因分析。尽管大多数内分泌疾病并非遗传性疾病，但是从遗传性疾病的致病基因分析过程中所获得的对激素作用相关基因结构与功能的认识，将直接指导其他具有类似病理变化和临床表现的疾病的诊断和治疗。以下分别列举了肽类激素分子自身异常的原因（表 15-1）和靶细胞敏感性异常的原因（表 15-2）。

由于内分泌系统的复杂性，同一种内分泌系统疾病，虽然临床表现可能相同，但是引起这些临床表现的分子机制却可能不尽一致。

表 15-1　激素分子结构和含量异常可能原因

原　　因	可能引起的表型变化
激素自身的基因突变	mRNA 异常：mRNA 水平低于或高于正常
	mRNA 分子大小异常
基因编码区突变	mRNA 稳定性降低或升高
基因调控区突变	
	蛋白质改变：激素活性升高或降低
	激素分子大小改变
	激素细胞内定位改变
	激素稳定性增强或下降
特异性转录调控因子的异常	mRNA 水平低于或高于正常
	激素水平低于或高于正常
激素分泌相关分子的异常	激素细胞内定位改变、激素分泌障碍

表 15-2　激素靶细胞敏感性异常种类

相关分子种类	功 能 变 化
受体	
	受体数目的增加、减少或缺如
	受体结构异常导致亲和力下降或增强
细胞内信号转导分子	
各种蛋白激酶、蛋白磷酸酶、接头蛋白、GTP 结合蛋白、转录因子、细胞骨架蛋白等	活性升高或降低、细胞内定位改变、相互作用分子变化、调节方式变化、稳定性增强或下降
细胞效应相关分子	
能量代谢酶类、细胞周期调节因子、核酸和蛋白质合成体系分子、细胞存活与凋亡调节和效应分子、离子通道	活性升高或降低、细胞内定位改变、相互作用分子变化、调节方式变化、稳定性增强或下降

第二节　内分泌代谢疾病的分子机制举例

内分泌代谢疾病有的属于遗传病，目前已经发现了相当多的致病基因。有的是偶发的、非遗传的先天性异常；有的是具有遗传易感性的多基因决定的疾病；有的是环境、营养等造成。尽管近年来，随着人类基因组测序工作的完成，对于许多内分泌疾病的分子机制有了较多的认识，但是对大部分疾病的致病基因的分析工作仍在进行之中。这里首先介绍侏儒症这一致病基因相对明确的疾病，然后再介绍糖尿病这一内分泌系统中最常见、分子机制较复杂疾病的有关进展。

一、侏儒症

侏儒症（dwarfism）病人大部分是由于生长激素（growth hormone，GH）缺乏所致，少部分是由于靶细胞对 GH 不敏感所致。根据侏儒症的病变特征和对一些家系的致病基因分析，现已累计报道有 522 种不同类型的侏儒症存在（见 http//www. dwarfism. org/medical/types. php）。涉及的分子病理机制从下丘脑-垂体，直至与生长相关的 GH 靶细胞，都有一些异常的基因或功能的异常被发现。表 15-3 列举了部分已经确认的部分分子异常。

表 15-3　侏儒症部分代表性分子异常

类　型	致病基因	突变类型	突变效应	遗传方式
Ⅰ型垂体性侏儒症	GH	GH 基因缺失、GH 基因剪接位点突变	GH 水平低下	散发或常染色体隐性
Ⅱ型垂体性侏儒症	GH 受体	第 8 外显子缺失、R43X 点突变、其他点突变	细胞对 GH 无反应、血中 IGF 水平降低	常染色体隐性、常染色体显性
Ⅲ型垂体性侏儒症	PIT1 转录因子	点突变、移码突变、基因剪接位点突变	联合垂体激素低下	常染色体隐性、常染色体显性
Ⅳ型垂体性侏儒症	GH	不明，可能在 17 号染色体有突变	形成多聚体 GH	散发

二、糖尿病的分子机制

糖尿病是一类以高血糖为特征的内分泌代谢疾病，在病因学、病理学和临床诊断治疗都存在极大的异质性。1 型糖尿病表现为血中胰岛素绝对水平低下，主要病变是胰岛 B 细胞在自身免疫机制下发生损伤和凋亡，胰岛素治疗有效；2 型糖尿病是由于肌和脂肪组织对胰岛素反应能力低下所致。尽管糖尿病发生的原因十分复杂，但是基于分子生物学的进步、动物模型和遗传性糖尿病的致病基因分析技术和理论的发展，近年来对于糖尿病发生和发展的分子机制有了较多的认识。由于 1 型糖尿病与自身免疫反应密切相关，故本书第十六章将对自身免疫反应的发生机制及其遗传易感性问题予以讨论。本章则主要从内分泌代谢调节角度讨论其有关的分子机制。

（一）胰岛素分泌的遗传缺陷

虽然大部分 1 型糖尿病并非遗传性疾病，但是对少数具有遗传性的糖尿病家系，即常染色体显性遗传的青年-成年发病型糖尿病（maturity-onset diabetes of the young，MODY）的致病基因分析，为认识胰岛素分泌异常的分子机制提供了机会。目前已经发现至少 6 个基因与胰岛素分泌异常有关，其中一个是编码葡萄糖激酶的基因，另外 5 个都是转录因子，包括 3 个肝细胞核因子（hepatocyte nuclear factor，HNF）、IPF-1 和胰岛 B 细胞 E-盒反式激活因子 2（B-cell E-box transactivator 2，BETA2），它们的缺陷和功能障碍见表 15-4。

表 15-4　胰岛素分泌缺陷相关基因

基因名称	基因产物的功能	突变效应	致 病 性
葡萄糖激酶基因（*MODY* 2）	催化 6-磷酸葡萄糖生成，作为胰岛 B 细胞的葡萄糖感受器	胰岛 B 细胞葡萄糖诱导的胰岛素分泌减少、肝糖原储存减少	纯合子发生先天性糖尿病，杂合子发生 MODY
HNF-4（*MODY* 1） *HNF*-1（*MODY* 3） *HNF*-1（*MODY* 5）	转录因子，共同控制胰岛素、糖转运蛋白、糖酵解酶、线粒体酶的合成	胰岛 B 细胞中葡萄糖和其他信号诱导的胰岛素分泌减少	杂合子发生 MODY
IPF-1（*MODY* 4）	转录因子，与胰腺发育有关、控制胰岛素、糖转运蛋白、葡萄糖激酶的合成	胰岛 B 细胞中葡萄糖诱导的胰岛素分泌减少	纯合子发生胰腺发育障碍和先天性糖尿病，杂合子发生 MODY
BETA2（*MODY* 6）	转录因子，与胰腺发育有关、控制胰岛素合成	血中胰岛素水平降低	偶发 MODY

除了上述 6 种基因的产物影响胰岛素的合成与分泌以外，其他的胰岛 B 细胞转录因子的突变也可能与 MODY 发生有关。此外，线粒体基因组的突变也会导致胰岛素分泌障碍。

（二）胰岛素受体介导信号转导的异常

2 型糖尿病的重要特征是胰岛素抵抗效应，主要的糖代谢相关组织对胰岛素的反应能力下降。理解胰岛素不敏感的分子机制对于 2 型糖尿病的诊断和治疗都具有重要意义。要认识胰岛素不敏感的分子机制，必须首先阐明正常情况下的胰岛素受体介导的信号转导通路。

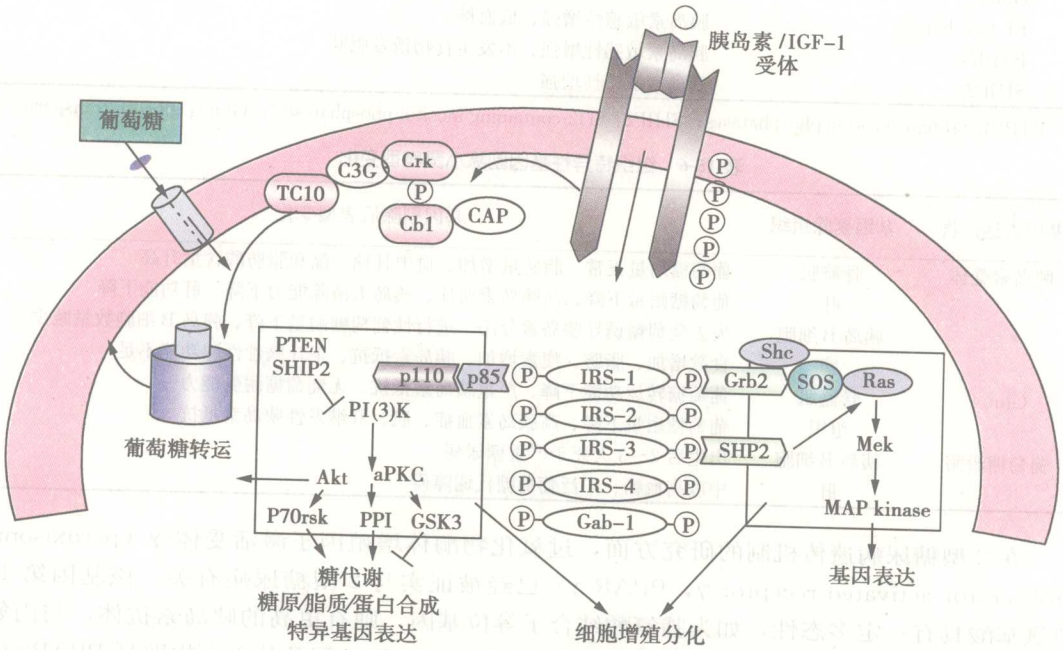


图 15-4　胰岛素受体介导的信号转导

图中缩写：TC10：Ras-like protein TC10；C3G：CRK SH3-binding protein；Cbl：v-oncogene from Cas-Br-M virus；CAP：Cbl-associated protein；CRK：transforming oncogene v-crk of avian sarcoma virus CT10；GSK：glycogen synthase kinase；SHIP：SH2-containing inositol phosphatase；Rsk：ribosomal S6 kinase；Gab：Grab2-associated binding protein；Akt：PKB

胰岛素受体介导的信号转导通路及其对细胞功能的影响如图 15-4 所示。葡萄糖的转运、葡萄糖代谢、细胞增殖与分化是胰岛素作用后的主要效应。在胰岛素信号转导通路中，胰岛素受体底物-1（insulin receptor substrate-1，IRS-1）是重要的信号衔接蛋白，PI3K-PKB、MAPK、Cbl 等通路是几个关键通路。表 15-5 列举了上述通路相关的基因剔除小鼠模型的表型变化。这些变化的发现对于确认糖尿病的发病机制，以及治疗药物发现有重要帮助。

在组织特异性基因剔除小鼠观察到的代谢异常（表 15-6），更为准确地反映了胰岛素受体介导的信号转导通路中关键分子的作用。

（三）糖尿病的遗传易感问题

从分子生物学角度，糖尿病不是一种单纯的病理过程，而是一组具有遗传异质性的疾病，其共同表现为葡萄糖耐量降低，即不同遗传背景，不同遗传与环境相互作用方式，但是有共同的表型。2 型糖尿病的遗传倾向大于 1 型糖尿病。

表 15-5　胰岛素受体信号通路相关分子基因剔除小鼠表型变化

基因表达产物	基因剔除表型变化
胰岛素受体	宫内正常生长，出生后 3～7 天死于糖尿病性酮血症
IGF 受体	宫内及出生后生长迟缓，血糖稳定
IRS	胰岛素抵抗、葡萄糖耐量降低、生长迟缓
IRS-2	胰岛素抵抗、胰岛 B 细胞发育障碍、2 型糖尿病
IRS-3	正常
IRS-4	正常
Akt 2	肝、肌胰岛素抵抗
Glut4	心肌肥大、心功衰竭、糖耐量正常
PI（3）K-p85	胰岛素敏感性增强、低血糖
PTP1B	胰岛素敏感性增强、不发生食物诱发肥胖
SHIP2	胰岛素敏感性增强

PTPIB：protein-tyrosine phosphatases；SHIP2：SH2-containing inositol phosphatase 2；Glut4：glucose transporter 4

表 15-6　组织特异性基因剔除小鼠表型变化

基因表达产物	基因剔除组织	基因剔除后表型变化
胰岛素受体	骨骼肌	葡萄糖耐量正常、脂肪量增加、血中甘油二酯和脂肪酸含量升高
	肝	葡萄糖耐量下降、高胰岛素血症、胰岛素清除能力下降、肝功能下降
	胰岛 B 细胞	失去葡萄糖诱导胰岛素分泌、进行性葡萄糖耐量下降、胰岛 B 细胞数量减少
	脑	食欲增加、脂肪、瘦素增加、胰岛素抵抗、下丘脑性性腺功能不足
Glut 4	骨骼肌	葡萄糖转运功能下降、严重胰岛素抵抗、无葡萄糖耐受能力
	组织	葡萄糖耐量下降、高胰岛素血症、肌、肝继发性胰岛素抵抗
葡萄糖激酶	胰岛 B 细胞	出生后 2～3 天死于严重糖尿病
	肝	中度高血糖、广泛葡萄糖代谢障碍

在 2 型糖尿病遗传机制的研究方面，过氧化物酶体增殖因子激活受体-γ（peroxisome proliferator-activated receptor-γ，PPAR-γ）已经被证实与 2 型糖尿病有关。该基因第 12 位氨基酸具有一定多态性，如为脯氨酸纯合子等位基因，则有更高的胰岛素抗体，与丙氨酸等位基因相比，发生糖尿病的危险性提高 12 倍。噻唑烷二酮及其衍生物即是 PPAPγ 的激动剂，可以提高机体对胰岛素的敏感性。

无论是 1 型还是 2 型糖尿病，都存在广泛的遗传异质性。由此提出了一个重要问题，如果糖尿病的易感性是遗传所决定的话，为什么在进化中没有被淘汰？Neel 提出，2 型糖尿病表型的遗传背景使得对食物的利用效率增加，在早期人类经常面临饥饿威胁时，无疑这样的遗传背景有着极大的生存优势，被自然选择所推崇。到了现代社会，可为人们提供的营养已经过剩，这样的遗传背景才会导致糖尿病和肥胖。这一学说解释了为什么随着现代化和都市化，糖尿病的发病率大大增加的一个原因。

（药立波）

第十六章　免疫系统疾病的分子机制

关键知识点

- 免疫球蛋白、B淋巴细胞受体、T淋巴细胞受体及人类白细胞抗原都属于重要的免疫识别分子。
- 抗体的基因表达、结构及功能异常引起B细胞免疫缺陷病；TCR及信号转导异常引起T细胞免疫缺陷病。重组酶RAG是Ig和TCR基因重排的关键酶，其基因突变可引起重症联合免疫缺陷病。
- 自身免疫病的发病机制是自身免疫耐受失调。自身抗体和自身反应性T细胞针对自身抗原产生免疫应答并造成病理性损伤。HLA的结构与功能特点决定其为器官特异性和系统性自身免疫病的易感因素。

关键词：免疫缺陷病；自身免疫病；免疫球蛋白；抗体；B细胞受体；T细胞受体；人类白细胞抗原；自身免疫病易感基因

免疫系统是机体清除外来侵入的病原体、监视和维持体内免疫应答稳态的主力军。B淋巴细胞、T淋巴细胞、巨噬细胞、树突状细胞及自然杀伤细胞等免疫细胞是发挥免疫作用的功能单元；免疫球蛋白（immunoglobulin，Ig）、B细胞受体（B cell receptor，BCR）、T细胞受体（T cell receptor，TCR）、人类白细胞抗原（human leukocyte antigen，HLA）、细胞因子及补体等则是执行免疫功能的分子基础。免疫细胞在发育成熟、激活、增殖和分化过程中，伴有免疫分子表达的阶段特异性变化，与机体免疫应答的需要相适应，并且也是反映机体免疫功能正常与否的重要指标。免疫系统疾病发生的分子机制与遗传、理化因素、感染等引起的免疫分子表达、结构与功能异常造成的免疫细胞功能障碍及免疫应答失调密切相关。免疫系统疾病主要分为免疫缺陷病、自身免疫病、超敏反应和淋巴增殖性疾病四类。本章将概要性介绍免疫缺陷病和自身免疫病的分子病理机制。

第一节　免疫缺陷病的分子机制

免疫缺陷病是免疫功能低下或缺失所引起的，以对各种微生物的反复感染为特征的临床综合征的总称。免疫缺陷病包括原发性免疫缺陷病（primary immuno-deficiency diseases，PID）和继发性免疫缺陷病（secondary immuno-deficiency diseases，SID）。PID源自于免疫器官、细胞的遗传和发育障碍或免疫分子的基因缺陷。SID常由恶性肿瘤、HIV感染及应用免疫抑制剂等引起。本节仅介绍PID。PID可通过常染色体隐性（86%）、常染色体显性（4%）及X连锁（10%）方式遗传。目前已报道有150余种PID，发现了100多个致病基因，并从中鉴定出4 500种基因突变类型。PID分为B细胞免疫缺陷、T细胞免疫缺陷、联合免疫缺陷、巨噬细胞缺陷和补体缺陷等不同种类。这里简介B细胞免疫缺陷病、T细胞免疫缺陷病和联合免疫缺陷病的分子机制，其他类型的PID及相关基因和蛋白质结构、遗传及临床等资料等可参考相关数据库（http：//bioinf. uta. fi/IDbases）。

一、抗体基因的结构与表达异常引起 B 细胞免疫缺陷病

B 细胞在体液免疫应答中起关键作用。成熟的 B 细胞产生分泌型抗体（Ig）及膜型抗体（mIg，即 BCR）。抗体基因表达、结构及功能异常引起 B 细胞功能障碍，并导致 B 细胞免疫缺陷病。

（一）Ig 的分子结构及主要功能特点

1. Ig 的共同结构特点　各种 Ig 单体都具有由两条重链（H）和两条轻链（L）构成的“Y”型基本结构。Ig 通过链间二硫键连接重链和轻链，借助链内二硫键构成特征性的免疫球蛋白结构域（Ig domain）。除 Ig 外，还有多种免疫分子，如 BCR、TCR、CD3、CD4、CD8、Igα、Igβ、Fc 受体（FcR）等也含有免疫球蛋白结构域。Ig 的主要功能区是由重链和轻链的可变区（V_H 和 V_L）构成的抗原结合部位，及重链恒定区（C_H）的补体结合部位和巨噬细胞等细胞表面的 FcR 结合部位。

2. Ig 的基因组构成和基因重排　编码 Ig 重链和轻链（κ、λ）的基因分别位于染色体 14q32、2p12 和 22q11，基因全长 1.3 Mb、2.3 Mb 和 880 kbp。B 淋巴细胞在分化成熟过程中发生 Ig 基因的重排，图 16-1 显示了 IgM 重链的基因组构成及基因重排和表达过程。

在基因重排的识别、剪接和连接过程中，重组酶 RAG1/RAG2 起关键作用。RAG1 和 RAG2 分别由重组激活基因 1（recombinant activating gene 1）和重组激活基因 2 编码。

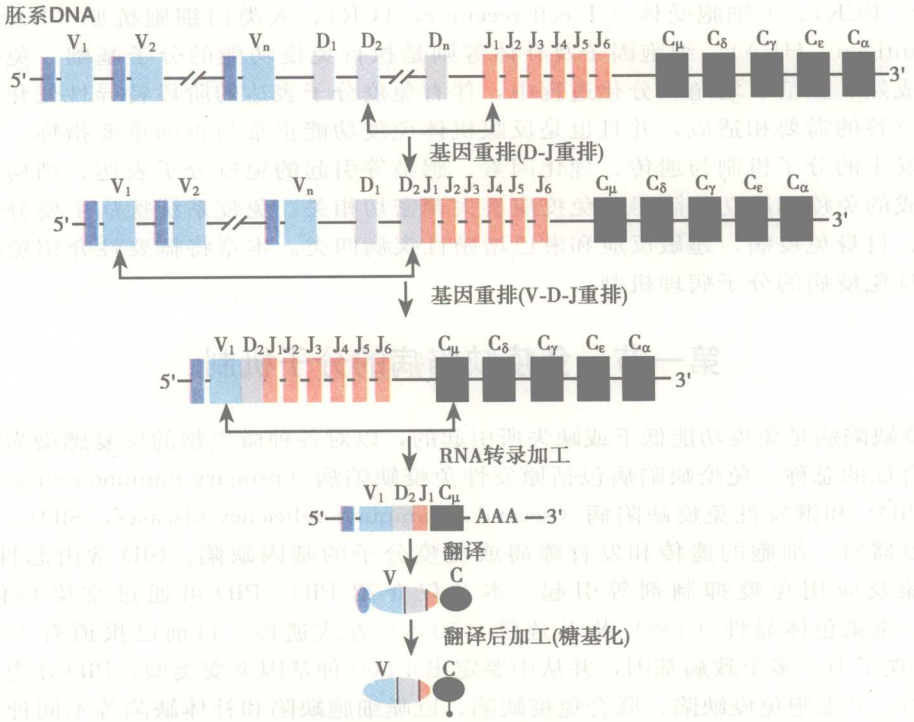


图 16-1　IgM 重链的基因组构成及基因重排和表达

IgM 重链基因组（胚系 DNA）中有 4 种基因片段（V，D，J，C），各有 30～60 种功能型基因片段。V，D，J 编码 VH，$C\mu$ 编码 CH。D-J 基因重排形成 D_2-J_1 重组片段，V-D-J 基因重排形成 V_1-D_2-J_1 重组片段。V_1-D_2-J_1 与下游的 $C\mu$ 连接形成完整的初始 mRNA，经转录加工成为成熟的 mRNA，并作为模板指导 IgM 重链合成。翻译后的产物经糖基化等修饰变成有功能的 IgM 重链

RAG1/RAG2 为 B 细胞和 T 细胞特异性表达的酶，对其发育成熟、激活和分化中的 Ig 和 TCR 合成非常重要。

3. 不同类型 Ig 的免疫功能　IgG 占血清 Ig 总量的 80%，是唯一能穿过胎盘的抗体，具有中和作用、激活补体和调理作用，对新生儿的抗感染能力起重要作用；分泌型 IgA 存在于乳汁、唾液、泪液及呼吸道、消化道和泌尿生殖道等黏膜表面的分泌液中，在局部免疫应答中起重要作用；分泌型 IgM 是初次体液免疫应答中最早产生的抗体，其补体结合和激活能力作用强，膜型 IgM（mIgM）在早期 BCR 即开始表达；膜型 IgD（mIgD）参与成熟的 BCR 的构成，是 B 细胞成熟和分化的重要标志之一；IgE 的血清含量极低，其 Fc 段可与肥大细胞和嗜碱细胞等的 FcR 结合，具有抗病原体感染作用。各种 Ig 的合成水平、结构、性质、在体内的分布及在体液免疫应答中的作用不同，并直接影响体液免疫功能。

（二）BCR 的结构、表达特点及其主要功能

1. BCR 的结构特点　BCR 是由膜型 Ig（mIgD、mIgM）和 Igα/Igβ 构成的复合体（图 16-2）。膜型 Ig 主要见于非激活状态下的 B 细胞，当 B 细胞在抗原刺激下发生增殖和分化时，则大量产生分泌型的 Ig。未成熟的 B 细胞仅表达 mIgM，而成熟的 B 淋巴细胞共同表达 mIgD 和 mIgM。Igα 和 Igβ 分子的胞浆部分结构中含有免疫受体酪氨酸激活模体（immuoreceptor tyrosine-based activation motif，ITAM），其序列（Tyr-X-X-Leu/Ile-X…$X_{(6\sim9)}$，X 代表任意氨基酸）中含有激活后可被磷酸化的酪氨酸残基。Igα 和 Igβ 通过可传递抗原结合 BCR 后所产生的激活 B 细胞激活的特异刺激信号。

2. BCR 在 B 细胞发育成熟过程中的特异表达　B 细胞发育成熟过程中，经历了造血干细胞→淋巴干细胞→祖 B 细胞→前 B 细胞→未成熟 B 细胞→成熟 B 细胞→浆细胞阶段，同时伴有 BCR 组成性的阶段特异变化。例如，在前 B 细胞期出现膜型 IgM 前体（Igμ）和替代轻链（λ5）；成熟 B 细胞期共同表达 mIgM 和 mIgD；抗原刺激使成熟的 B 细胞激活并发生 Ig 类别转换，由最先合成 IgM，转变为产生 IgG、IgA 或 IgE 的浆细胞。这种特异表达反映了 B 细胞的发育成熟和功能状态。

3. BCR 的主要功能　BCR 的主要功能是与抗原特异的识别和结合及向细胞内传递激活信号，其介导的信号转导途径对 B 细胞的发育成熟、激活和分化等起重要作用。

B 细胞的激活依赖于 BCR 与抗原特异的识别和结合，并引起细胞内多种分子的相互作用和蛋白酪氨酸激酶的级联激活（图 16-2）。3 种 PTK（Src、Syk 和 Tec）参与 B 细胞激活。PKC、MAPK 等信号转导途径的激活启动 B 细胞相关分子的表达，从而引起体液免疫功能变化。此种 BCR 与抗原的结合是 B 细胞激活的第一信号。B 细胞激活尚需要激活的 T 细胞膜上的 CD40 配体（CD40L）和 B 细胞膜上的 CD40 分子间的相互作用（第二信号或称协同刺激信号）。

（三）B 细胞免疫缺陷病

1. B 细胞免疫缺陷病的分子机制　B 细胞发育成熟、激活和分化等各阶段免疫分子的基因表达和功能异常，特别是参与细胞识别和信号转导的分子都可引起 B 细胞免疫缺陷病。例如，促淋巴细胞成熟的转录因子、细胞因子及其受体（IL-7、IL-7R）等表达的异常，可阻止造血干细胞向淋巴干细胞的分化，并使 B 细胞和 T 细胞的发育和功能均受影响；与 B 细胞成熟及激活信号通路相关的各种分子的基因缺陷，均可引起 B 细胞免疫缺陷病。

2. B 细胞免疫缺陷病的特征　B 细胞免疫缺陷病占 PID 的 50%～70%，其共同特点是体液免疫功能低下，多见于婴幼儿，常伴有血清 Ig 含量低或类型异常、淋巴组织和血液中 B 细胞减少等，临床表现为反复发生的细菌感染。B 细胞免疫缺陷病多数有遗传倾向。体液免疫功能的检测包括：血清 Ig 含量和种类测定，外周血白细胞总数和淋巴细胞

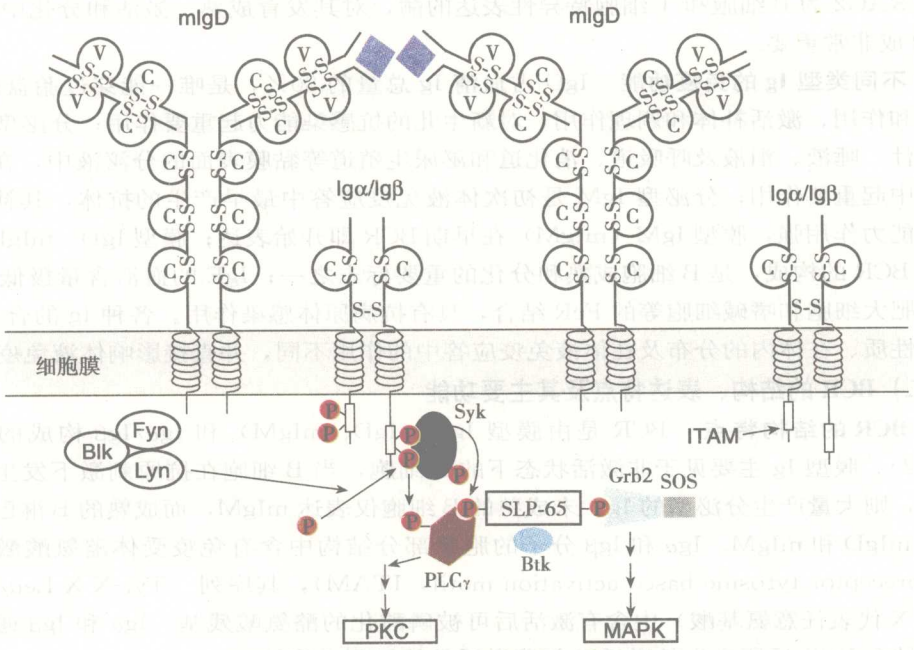


图 16-2　BCR 结构及 B 细胞激活的信号转导途径

BCR 由 mIgD 和 Igα/Igβ 组成。Igα/Igβ 胞浆内的部分有 ITAM。抗原结合 mIgD 使 mIgD 分子交联，引起 Igα/Igβ 中 ITAM 的酪氨酸磷酸化。Src 家族成员（Blk、Fyn、Lyn）介导 ITAM 的磷酸化。ITAM 磷酸化进一步募集并激活 Syk。磷酸化的 Syk 可连接适配蛋白 SLP-65。SLP-65 使 Btk（Bruton tyrosine kinase）的 Tyr551 磷酸化及 Tyr223 的自身磷酸化。SLP-65 结合 PLC（磷脂酶 C）γ 及 Grb2 等多种分子激活下游的 PKC 和 MAPK 等信号转导途径

分类计数，及 B 细胞抗体产生能力、活化和增殖功能等，均可辅助诊断 B 细胞免疫缺陷病，并为疾病的治疗和预后判断提供依据。

3. B 细胞免疫缺陷病举例　与 Ig 有关的 B 细胞免疫缺陷病有 10 余种，主要分为两类，一类是无丙种球蛋白血症（X 连锁无丙种球蛋白血症、IgM 重链缺陷、Igα 链缺陷、λ5 替代轻链缺陷、κ 轻链缺陷、X 连锁低丙种球蛋白血症伴生长激素缺陷等）；另一类是选择性 Ig 缺陷病（选择性 IgA 缺陷、选择性 IgG 亚型缺陷、选择性 IgA 亚型缺陷、选择性 IgE 缺陷及 X 连锁高 IgM 血症等）。

（1）X 连锁无丙种球蛋白血症(X-linked agammaglobulinemia，XLA)：XLA 又称为 Bruton 病，是由位于 Xq21.3-22 上的 *BTK* 基因突变和缺失所致。*BTK* 基因全长37.5 kb，含有 19 个外显子，编码 77 kDa 的酶蛋白。目前，已经在 973 个 XLA 家族(1 111 例病人)中发现了 600 余种基因突变。基因突变使酶分子空间结构的破坏及酶活性丧失。Btk 是第一个被发现与遗传病相关的蛋白酪氨酸激酶，它在 B 细胞发育成熟和分化的信号转导中起重要作用。XLA 患者的 B 细胞发育受阻于祖 B 细胞向前 B 细胞分化的阶段，导致外周血中只有极少数或缺乏成熟的 B 细胞。血清 IgG 及 IgM 和 IgA 含量均显著降低。

（2）选择性 IgA 缺陷病：选择性 IgA 缺陷病主要由 Ig 类别转换障碍引起。Ig 的 C_H 基因重排异常，使 B 细胞不能转变成分泌 IgA 的浆细胞，并导致 IgA 的合成和分泌水平减少。患者血清中的 IgA 浓度低，而 IgG 和 IgM 含量可正常或偏高。

（3）X 连锁高 IgM 血症（X-linked hyper-IgM syndrome，XHIM）：XHIM 也是由 Ig 类别转换障碍引起。T 细胞的 CD40L 基因位于染色体 Xq26-27.2，该基因突变使 CD40L 不能与 B 细胞的 CD40 结合，无法形成 B 细胞激活和分化的第二信号，导致 B 细胞的 Ig

类别转换间接的受到影响。患者血液及淋巴组织中分泌 IgM 的 B 细胞增加，血清中 IgG、IgA 和 IgE 含量极低或检测不到，而 IgM 水平异常升高，并且此种 IgM 主要是针对红细胞和血小板等的自身抗体。患者还常伴有细胞免疫功能的缺陷。

二、TCR 基因结构与表达异常引起 T 细胞免疫缺陷病

T 细胞在细胞免疫应答中起关键作用。TCR 参与 T 细胞的发育成熟、激活和分化过程，其基因表达、结构及功能的异常引起 T 细胞功能障碍，并导致 T 细胞免疫缺陷病。

（一）TCR 结构特点及其介导的 T 细胞激活信号转导途径

1. TCR 的结构特点　TCR 包括 αβTCR（主要类型）和 CD3 两部分（图 16-3）。αβTCR 的 α 链和 β 链相似于 Ig 的重链和轻链，决定 TCR 的特异性；二者在合成中也需要在 RAG1/RAG2 作用下发生基因重排，这也是产生 TCR 多样化的原因。CD3 的作用与 BCR 的 Igα/Igβ 相同，参与抗原刺激信号向细胞内的传递。T 细胞膜表面还有其他分子，参与细胞和分子间的相互作用及辅助信号传递，如 CD4 和 CD8 可分别与 HLA-Ⅱ类分子和 HLA-Ⅰ类分子结合。

2. TCR 介导的信号转导途径　TCR 所介导的 T 细胞激活信号转导途径，与 B 细胞激活的信号转导途径机制和过程相似（图 16-3）。例如，ZAP-70（zeta associated protein-70）与 Syk 属同一家族，在 T 细胞信号转导途径起重要作用。TCR 结合抗原后可激活 T 细胞内的信号传导途径，诱导 T 细胞激活相关的免疫分子表达及细胞增殖和分化，产生细胞免

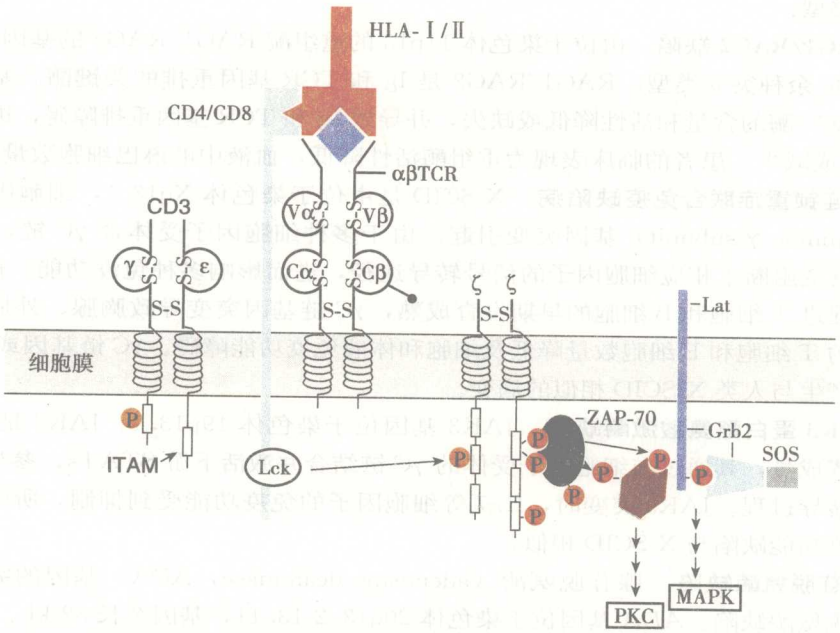


图 16-3　TCR/CD3 复合体结构及 T 细胞激活的信号转导途径

TCR/CD3 复合体包括 αβTCR 和 CD3。αβTCR 的 α 和 β 链分别由 Vα 和 Cα 及 Vβ 和 Cβ 组成。Vα 和 Vβ 为 T 细胞与抗原肽/HLA 复合物结合部位。CD3 多肽链含 γε 和 ζζ 等二聚体，其胞浆内末端有 ITAM。T 细胞激活信号转导途径开始于抗原肽/HLA（HLA-Ⅰ类/HLA-Ⅱ类）复合物与 TCR/CD3 复合物和 CD4/CD8 的结合。CD4/CD8 的胞浆内部分与 PTK 家族成员 Lck（lymphocyte-specific protein tyrosine kinase）相连并激活 Lck，Lck 使二聚体 ζζ 的 ITAM 的酪氨酸磷酸化。ZAP-70 与磷酸化酪氨酸结合被激活及自身磷酸化激活。作为适配分子，ZAP-70 激活磷酸化其他分子，如 Lat（linker of activated T cells）、Grb2 及 PLCγ 等，激活 PKC、MAPK 等信号转导途径

疫效应。

（二）T细胞免疫缺陷病

1. T细胞免疫缺陷病的分子机制　编码 Pre-Tα（TCRα 前体）、TCRβ、CD3（CD3γ、CD3ε）、CD4、CD8、Lck、ZAP-70 等蛋白的基因突变和缺失均可引起细胞免疫功能缺陷，并导致 T 细胞免疫缺陷病。例如，*ZAP-70* 基因缺陷导致患者胸腺和外周免疫器官的 $CD8^+$ T细胞缺失，造成细胞免疫功能低下。

2. T细胞免疫缺陷病的特征　T 细胞免疫缺陷病包括 TCR 等膜分子和信号转导分子的基因突变所引起的 T 细胞免疫缺陷病，及先天性胸腺发育不全（DiGeorge 综合征）等。T 细胞免疫缺陷病的共同特点是细胞免疫功能低下，患者易受病毒、真菌及细胞内病原体的感染，并常伴有自身免疫病和恶性肿瘤。细胞免疫功能的检测包括：外周血白细胞总数和淋巴细胞分类计数、细胞增殖试验、细胞毒试验及细胞因子产生能力等，有助于 T 细胞免疫缺陷病的诊断、治疗和预后判断。

三、T细胞和B细胞异常引起严重联合免疫缺陷病

重症联合免疫缺陷病（severe combined immunodeficiency diseases，SCID）是联合免疫缺陷病的一种，特点是细胞免疫功能和体液免疫功能均严重低下或缺失。SCID 可由多种基因突变所导致的 T 细胞和 B 细胞的发育成熟、激活、分化及代谢障碍等引起，多见于婴幼儿，临床表现为 T 细胞和 B 细胞数目减少，并反复发生严重感染等。SCID 主要有以下几种类型：

1. RAG1/RAG2 缺陷　由位于染色体 11p13 的重组酶 RAG1/RAG2 的基因突变引起，已经发现 30 余种突变类型。RAG1/RAG2 是 Ig 和 TCR 基因重排的关键酶。基因突变使 RAG1/RAG2 酶的含量和活性降低或缺失，并导致 Ig 和 TCR 基因重排障碍，进而造成 Ig 和 TCR 合成减少。患者的临床表现为重组酶活性降低，血液中的淋巴细胞数量减少。

2. X 连锁重症联合免疫缺陷病　X-SCID 是由位于染色体 Xq13.1，细胞因子受体的 γC 链（common γ subunit）基因突变引起。由于多种细胞因子受体含 γC 链，因此，γC 链基因的突变阻断了相应细胞因子的信号转导途径，进而影响多种免疫功能。例如，IL-7 的作用是促进 T 细胞和 B 细胞的早期发育成熟，γC 链基因突变导致胸腺、外周免疫器官和血液中的 T 细胞和 B 细胞数量降低及细胞和体液免疫功能障碍。γC 链基因或 IL-7 基因敲除小鼠产生与人类 X-SCID 相似的病变。

3. JAK3 蛋白酪氨酸激酶缺陷　JAK3 基因位于染色体 19p13.1。JAK3 是 JAK 酪氨酸激酶家族成员，其通过与细胞因子受体的 γC 链结合并激活下游 STATs，参与了细胞因子的信号转导过程。JAK3 突变时，IL-7 等细胞因子的免疫功能受到抑制，所引起的细胞和体液免疫功能缺陷与 X-SCID 相似。

4. 腺苷脱氨酶缺陷　腺苷脱氨酶（adenosine deaminase，ADA）基因的突变和缺失引起腺苷脱氨酶缺陷。*ADA* 基因位于染色体 20q13.2-13.11，基因全长 32 kb。*ADA* 催化腺苷和脱氧腺苷脱氨，因此是嘌呤核苷酸分解代谢过程中的一个关键酶。*ADA* 基因突变使 ADA 的含量和活性降低或缺失，导致细胞内肌苷和脱氧肌苷及继发性升高的 dATP 蓄积，后者抑制核糖核苷酸还原酶而阻断 DNA 合成，进而影响细胞的代谢和增殖。发育中的淋巴细胞中降解 dATP 的 5′-核苷酸酶活性低。因此，淋巴细胞中的 DNA 合成受到明显抑制，患者 ADA 酶活性降低或检测不到，T 细胞和 B 细胞数目减少。对产前诊断可采用培养的微绒毛细胞或羊膜细胞进行酶活性检测和基因突变分析等。

5. 嘌呤核苷磷酸化酶缺陷　嘌呤核苷磷酸化酶（purine nucleoside phosphorylase，

PNP）基因定位于染色体 14q13.1。PNP 也是一个参与嘌呤核苷酸分解代谢的酶，其作用是使肌苷和脱氧肌苷转变成次黄嘌呤，及鸟苷和脱氧鸟苷转变成鸟嘌呤。*PNP* 基因突变导致的含量和活性降低或缺失，使细胞内鸟苷和脱氧鸟苷蓄积及继发性 dGTP 升高。*PNP* 基因突变主要是影响 T 细胞的代谢和功能，而对 B 细胞的影响相对较轻。临床表现类似于其他类型的 SCID。

免疫缺陷病的主要治疗原则是增强机体免疫功能，包括从基因、蛋白和细胞水平上纠正和替代缺陷的免疫分子和细胞等。对不同类型的 PID，可采取有针对性地治疗方法。例如，应用丙种球蛋白及基因工程抗体等可改善抗体异常所致的免疫缺陷病；同种异体胸腺组织移植可治疗先天性胸腺发育不全；ADA 缺陷的治疗包括基因治疗、酶替代治疗和造血干细胞移植，它也是世界上第一个进行基因治疗的疾病。另外，预防和治疗感染也是免疫缺陷病的一个常规治疗。

第二节　自身免疫病的分子机制

免疫系统具有识别“自己”与“非己”抗原的能力。在正常情况下，免疫系统通过多种调节方式建立并维持对自身抗原不产生或仅产生极微弱免疫应答的免疫耐受；但在某种异常条件下，免疫耐受失调，免疫系统针对自身组织、细胞或分子产生异常的免疫应答并造成病理性损伤，引起自身免疫病。目前已发现 40 余种自身免疫病，分为器官特异性自身免疫病和系统性自身免疫病。自身免疫病的发病机制极其复杂，目前尚不完全清楚，但一般认为是遗传、环境、感染及精神因素等共同作用的结果，其中遗传因素具有重要作用。人类白细胞抗原 HLA 作为主要组织相容性抗原系统，是自身免疫病的易感因素，与自身免疫病的发生和发展密切相关。本节简述 HLA 的结构与功能特点及其在自身免疫病中的发病机制。

一、HLA 的基因组成及表达特点决定其功能

（一）*HLA* 基因的特点及分析

1. *HLA* 基因结构　*HLA* 基因位于人 6 号染色体短臂（6p21），全长 4.68 Mb，分为 *HLA*-Ⅱ区、Ⅲ区和Ⅰ区。*HLA*-Ⅰ区和Ⅱ区分别编码 *HLA*-Ⅰ类分子和 *HLA*-Ⅱ类分子，每个区还包括多个基因座、基因位点（图 16-4）。例如，编码 *HLA*-Ⅱβ链的 *DRB* 基因位点共有有 16 个（*HLA-DRB*1～16）。*HLA* 基因在人群中具有丰富的多态性，如在 *DRB1*

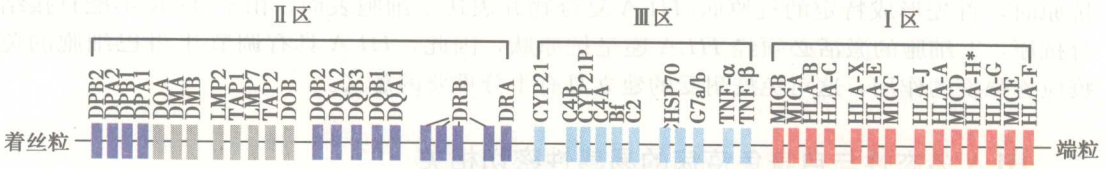


图 16-4　HLA 的基因组结构

HLA 基因组分为 HLA-Ⅱ区、Ⅲ区和Ⅰ区。HLA-Ⅰ区含编码经典的 HLA-Ⅰ类分子的基因座（HLA-A、HLA-B 和 HLA-C）和编码非经典的 HLA-Ⅰ类分子的基因座（HLA-E、HLA-F、HLA-G、HLA-H 及 MIC 等）。HLA-Ⅱ区（D 区），含编码经典的 HLA-Ⅱ类分子的基因座（DR、DP 和 DQ）。其中，DRA 和 DRB、DPA 和 DPB、DQA 和 DQB 编码 HLA-Ⅱ类分子的 HLA-Ⅱα和 HLA-Ⅱβ链。DRB 等有不同的基因位点。Ⅱ区还有编码非经典的 HLA-Ⅱ类分子的基因，如 DO 和 DM 及蛋白酶体（LMP2 和 LMP7）基因及抗原加工相关转运蛋白（TAP1 和 TAP2）基因。HLA-Ⅲ区有补体（C2、C4、B 因子）、热休克蛋白 70（HSP70）及肿瘤坏死因子（TNF-α、TNF-β）基因

中就有 500 多种等位基因，目前采用 *DRB1* * 010101、*DRB1* * 1613 等数字排列命名，星号后的数字两个为一组，分别代表编码不同种 DR 抗原的基因（同血清抗原）、基因亚型及亚型序列多态性。HLA 命名和序列等资料可参见数据库（http：//www. ebi. ac. uk/imgt/hla/）。

2. HLA 的基因分型　经典的 HLA 分型有血清学方法和细胞学方法。分子生物学技术的快速发展使依据 *HLA* 基因序列的 *HLA* 基因分型方法得到越来越广泛的应用，包括以 PCR 技术为基础的分型方法和 *HLA* 基因芯片方法等。*HLA* 基因分型方法具有高特异性和高敏感性，如临床用微量样品，即可对 *HLA* 的基因座、基因位点、等位基因及亚型进行分析。*HLA* 的基因分型有助于发现和鉴别与 *HLA* 基因直接相关的疾病类型和亚型。

3. *HLA* 基因的遗传特点　*HLA* 等位基因呈共显性遗传。各种 *HLA* 等位基因出现的频率不同，等位基因在同一染色体上的特定组合呈紧密连锁，构成个体的单倍型。群体中 *HLA* 的单倍型达 10^7 种以上，是目前发现的最具多态性的分子。*HLA* 的一些单倍型与自身免疫病的发生有关（表 16-1）。

表 16-1　*HLA* 与自身免疫病的相关性

器官特异性自身免疫病	相关 *HLA*	全身性自身免疫病	相关 *HLA*
胰岛素依赖性糖尿病	*DR3*、*DR4*、*DR3/4*	强直性脊柱炎	*B27*
桥本甲状腺炎	*DR3*、*DR5*、*DR8*、*A2*	多发性硬化	*DR2*
毒性甲状腺炎	*DR3*、*DR8*	系统性红斑狼疮	*DR3*
重症肌无力	*DR3*、*B8*、*A1*	类风湿关节炎	*DR4*
自身免疫性溶血性贫血	*B8*、*B10*	肺出血肾炎综合征	*DR2*
特发性血小板减少性紫癜	*DR2*、*B8*	干燥综合征	*DR3*
原发性慢性肾上腺皮质功能减退症	*DR3*	遗传性血色素症	*A3/B14*
原发性胆汁性肝硬化	*DR3*、*B8*		

4. *HLA* 基因的表达调控　*HLA*-Ⅰ类分子表达于各种有核细胞，而 *HLA*-Ⅱ类分子主要在成熟的 B 淋巴细胞、树突状细胞、巨噬细胞等免疫细胞中表达。*HLA* 的表达受多种因素的调控。感染、细胞因子及药物等可诱导 *HLA* 分子在一些细胞中异常表达，两者均可引起 *HLA* 相关的自身免疫病。

（二）HLA 的主要功能

HLA 的主要功能是通过分子间的相互作用，向 T 淋巴细胞递呈抗原并调节其功能。*HLA* 可递呈来源于病原体、机体自身成分并经抗原递呈细胞处理的抗原肽，甚至 *HLA* 分子本身。*HLA*-Ⅰ类分子和 *HLA*-Ⅱ类分子分别负载内源性和外源性抗原肽。*HLA* 递呈抗原时，首先形成特定的抗原肽/*HLA* 复合物并表达于细胞表面。由于 TCR 不能直接结合抗原，T 细胞的激活必须经 *HLA* 递呈抗原肽，因此，*HLA* 具有调节 T 淋巴细胞的免疫应答强度的作用，这对免疫耐受的建立具有十分重要的意义。

二、HLA 多态性与自身免疫病的易感性密切相关

目前普遍认为，不同个体在自身免疫病易感性上的差异，在很大程度上是由遗传因素所决定的。其中，又以 *HLA* 的多态性关联最为突出。*HLA* 多态性是自身免疫病的易感因素。

（一）HLA 与自身免疫病相关

HLA 是第一个被发现与自身免疫病相关（association）的免疫识别分子，而且与大多

数自身免疫病都有不同程度的关联。相对风险度（relative risk，RR）是分析 HLA 与疾病关联程度的指标。RR 为携带某特定 HLA 抗原（等位基因）者的患病率与非携带者患病率的比值。RR 值越大，表明该种 HLA 抗原（基因）与疾病的相关性越强。在自身免疫病中，HLA 多态性与强直性脊椎炎（ankylosing spondylitis，AS）的相关性最强（RR 值为 90）。另外，HLA 等位基因类别在不同地区人群中出现的频率不同；一种自身免疫病可能与几种 HLA 同时关联，不同类型的自身免疫病也可与同一种或几种 HLA 关联；HLA-Ⅰ类分子和 HLA-Ⅱ类分子中，以 HLA-Ⅱ类分子与自身免疫病的关联更常见。

（二）HLA 变异引起自身免疫病的机制

自身免疫病的发病机制有多种，例如，免疫隔离部位抗原的释放、非特异性多克隆 B 细胞增殖、自身反应性淋巴细胞凋亡受抑制及 HLA 异常等。与 HLA 有关的机制主要包括：

1. 免疫耐受失调　免疫耐受包括中枢免疫耐受和外周免疫耐受。中枢免疫耐受通过 T 细胞和 B 细胞的发育成熟过程实现。T 细胞在胸腺发育成熟中经历阳性选择和阴性选择。该过程中，TCR 与自身抗原肽/HLA 复合物适当的识别和结合（结合域值）至关重要。在胸腺皮质区，TCR 与自身抗原肽/HLA 复合物的识别和结合可避免 CD4 和 CD8 双阳性 T 细胞发生凋亡（阳性选择）；在胸腺髓质区，TCR 与自身抗原肽/HLA 复合物有高亲和力（超过结合域值）的 T 细胞发生凋亡，低亲和力的 T 细胞则继续发育成 $CD4^+$ 或 $CD8^+$ T 细胞（阴性选择）。B 细胞在骨髓的发育成熟过程中，自身反应性 B 细胞的清除机制与上述相似。如果中枢耐受不完善，自身反应性 T 细胞或 B 细胞进入外周免疫系统，则可通过外周耐受进一步被清除或活性受抑制。

HLA 负载抗原肽能力的变异可导致免疫耐受失调，产生大量针对自身抗原起免疫应答的自身反应性 T 细胞和 B 细胞，进而造成病理性损伤。

2. *HLA* 基因表达调控、结构和功能异常　*HLA* 基因表达、结构和功能异常可由多种原因引起。例如，正常情况下不表达或低表达 *HLA* 的细胞，但在某些病理条件下异常高表达。胰岛 B 细胞、肾上腺细胞、甲状腺上皮细胞及脑细胞几乎不表达 HLA-Ⅱ类分子，但在胰岛素依赖性糖尿病（insulin-dependent diabetes mellitus，IDDM）、原发性慢性肾上腺皮质机能减退症（Addison 病）、桥本甲状腺炎（Hashimoto thyroiditis，HT）及多发性硬化（multiple sclerosis，MS）等发病时则有过度表达；感染引起的细胞因子（IFN-γ、IL-10、TNF-α 等）释放增加可上调 *HLA* 表达；理化因素和药物作用等均可改变 HLA 的结构，并影响 HLA 递呈抗原的方式和程度。

3. HLA 参与分子模拟作用　病原体某些蛋白的氨基酸序列与宿主细胞的自身分子（如细胞膜受体，HLA）的氨基酸序列相同，使 HLA 递呈自身分子，并引起免疫系统在清除病原体过程中产生交叉免疫应答，造成细胞代谢和功能异常。

（三）自身免疫病的特征

自身免疫病患者常可检测到异常的自身抗体和自身反应性 T 细胞，自身免疫病多数有遗传倾向，患者易伴发其他类型的自身免疫病及肿瘤，病情反复发作，慢性迁延，病理性损伤主要由异常免疫应答引起的炎症为主。

（四）HLA 是器官特异性自身免疫病的易感因素

器官特异性自身免疫病是指免疫系统针对特定的组织、器官、细胞及其成分产生的异常免疫应答。胰岛素依赖性糖尿病（IDDM）是一种典型的器官特异性自身免疫病，为多因素、多基因所导致的复杂疾病。

1. IDDM 的病理机制　一般认为自身抗体、自身反应 T 细胞及细胞因子均参与了

IDDM的病理性损伤过程，以自身反应T细胞的作用为主。在遗传、感染等因素的共同作用下，最终使胰岛B细胞进展性破坏，胰岛素分泌不足，引起血糖升高等代谢疾病和功能障碍。IDDM患者的血清中可检测到多种自身抗体，如抗胰岛素抗体、抗谷氨酸脱羧酶（glutamic acid decarboxylase，GAD）抗体、抗酪氨酸磷酸酶抗体及抗胰岛细胞抗体等。抗胰岛素抗体（主要为IgG）是最常见的自身抗体，在活动期IDDM患者的阳性率较高，治疗后逐渐下降。IDDM患者胰岛B细胞的HLA-Ⅰ类分子和HLA-Ⅱ类分子表达升高，尤其是活动期的HLA-Ⅱ类分子升高的更明显。IDDM患者的IFN-γ升高，并可上调HLA-Ⅱ类分子的表达。

2. IDDM的易感基因　通过全基因组扫描筛查，目前已经确定的IDDM易感基因位点有18个（*IDDM* 1～18），分别位于不同的染色体上，包含900多个候选基因（http：//tldbase.org/page/Loci/display）。两个重要的IDDM易感基因位点为*IDDM1*和*IDDM2*。*IDDM1*基因位点含*HLA*基因区，包括编码*HLA-DMB*、*HLA-DRB1*、*TAP1*、*TAP2*等44个免疫分子的基因，占全部IDDM易感因素的40%以上。

IDDM的易感性与编码HLA-Ⅰ类分子和HLA-Ⅱ类分子的基因多态性相关。例如，患IDDM的白种人中，91%～95%携带*HLA-DR3*、*DR4*或*DR3/4*基因型，其中*DR3/4*基因型占40%～50%；而正常人群中*DR3/4*基因型仅为5%。对DQB1肽链一级结构的分析发现，若DQB1编码的基因产物出现Arg52则为IDDM易感性，而Asp57则与IDDM无相关性。HLA-A＊0201基因型与柯萨奇病毒B4感染后通过分子模拟作用产生的自身免疫应答有关。由于柯萨奇病毒B4的P2-C蛋白序列（氨基酸残基38～43）与谷氨酸脱羧酶（GAD）的氨基酸序列（氨基酸残基260～265）相同，因此，HLA-A＊0201递呈GAD，并使免疫系统针对胰岛B细胞产生免疫应答，导致胰岛B细胞破坏。人工合成的GAD寡肽在实验动物观察到与人类IDDM相似的病理性改变。HLA-A＊2402也与IDDM密切关联，主要是由于HLA-A＊2402基因产物所形成的抗原肽结合槽结构异常，不能与自身抗原结合并形成自身抗原肽/HLA-Ⅰ类复合物。患者病情进展快，胰岛B细胞几乎被全部破坏。

另外，*IDDM2*位于第11号染色体的胰岛素基因区，含有胰岛素基因上游5′端的5个可变数目串联重复序列（VNTR）及酪氨酸羟化酶基因，占全部IDDM易感因素的10%左右。

准确的定位易感基因位点并确定特异基因基序列，有助于从基因水平上对IDDM进行早期预防、诊断和治疗干预。

（五）HLA是系统性自身免疫病的易感因素

系统性自身免疫病是免疫系统针对分布广泛的自身抗原，产生异常免疫应答造成的病理性损伤，常累及皮肤、关节、肾等多个系统和器官。

1. *HLA*-B27与强直性脊柱炎　强直性脊柱炎（ankylosing spondylitis，AS）的临床表现以脊椎关节的慢性炎症为主，并可累及其他组织和器官。*HLA-B27*基因与AS的关联最强，正常人群中的*HLA-B27*仅占3%～7%，而AS患者90%以上携带*HLA-B27*。现已发现了39种*HLA-B27*亚型，*HLA-B*＊2705和*B*＊2702为白种人AS患者的主要类型，*B*＊2704和*B*＊2707为亚洲AS患者的主要类型。*HLA-B27*亚型之间的差异仅在于编码HLA-Ⅰα链的α1和α2结构域中有个别氨基酸残基不同，如与疾病相关的*HLA-B*＊2705和与疾病不相关的*B*＊2709，仅为多肽链上116位置上一个氨基酸残基的不同，在前者为Asp，而后者为His。

*HLA-B27*作为AS的易感基因与以下有关：①*HLA*-B27分子结构异常：α1结构域的Cys67形成同二聚体或发生折叠错误，不能结合并递呈抗原肽，从而影响T细胞的功能，

或变成细胞表面的“异己”抗原引起免疫应答；②*HLA*-B27 过度表达：病原体感染引起 IL-10 和 TNF-α 的产生和释放增加，使 *HLA*-B27 表达上调；③分子模拟：如克雷伯氏菌肺炎氮合成酶的部分氨基酸序列与 *HLA*-B27 的相同，*HLA*-B27 递呈病原体的蛋白分子的同时，将 *HLA*-B27 分子本身递呈给 $CD8^+$ T 细胞，引起异常免疫反应。

2. HLA 与系统性红斑狼疮　系统性红斑狼疮（systemic lupus erythematosus，SLE）是一种由多因素、多基因引起的，表现为多系统和器官病理性损伤的系统性自身免疫病。在 SLE 患者血清中可检测到多种自身抗体，如抗核抗体、抗双链 DNA 抗体、抗磷脂抗体、抗红细胞或血小板抗体等。各种抗体针对机体自身的不同成分，并出现于 SLE 病程的不同阶段。SLE 的病理性损伤主要由抗原抗体形成的免疫复合物在组织和血管壁沉积，所引起的炎症反应而造成。

目前，已经报道 100 余个基因参与 SLE 的发病过程。通过全基因组扫描发现了 SLE 基因易感区域（1q23、1q25-31、1q41-42、2q35-37、4p16-15.2、6p11-21、12q24 和 16q12-13）。进一步对这些热点区域进行的精细基因定位和位置候选基因分析，确定了 13 个与 SLE 有明确关联的基因，包括 C-反应蛋白（CRP）基因、HLA-DRB1、TNF-α 等。

位于 SLE 易感区 6p11-21 的 *HLA* 基因，也是最早被发现的 SLE 易感基因，与 SLE 关联性最强。不同地区 SLE 患者的 *HLA* 基因型不同，常见的类型为 *DR2* 和 *DR3*。白种人 SLE 患者 *HLA* 基因型为：*DRB1* * 1501（DR2）、*DQB1* * 0602，*DRB1* * 0301（DR3）、*DQB1* * 0201，*DRB1* * 0801（DR8）和 *DQB1* * 0402；亚洲 SLE 患者为 *DRB1* * 1501 或 *DRB1* * 1502（DR2）。*HLA* 基因型尚与患者的自身抗体种类及临床表现有关，若 *HLA-DQA* 基因编码产物含 Glu34，*DQB* 基因编码产物含 Leu26，则常伴有抗 Ro 抗体和抗 LA 抗体。其他 HLA-Ⅱ基因尚未发现与 SLE 相关的直接证据。

SLE 的发病机制复杂，除 *HLA* 外，其他基因的表达及调控异常也与 SLE 发生的机制有关。例如，携带 TNF-α-308A 等位基因者对自身免疫病的易感性增加。

（六）自身免疫病的实验动物模型

多种自身免疫病的自发性、诱导性或转基因实验动物模型，可用于人类相应自身免疫病分子机制研究及药物作用等的观察。例如，自发性糖尿病（NOD）小鼠，在出生后 3～6 周时出现抗胰岛 B 细胞抗体并且抗体滴度持续增加，待发育至 15 周时，90%的胰岛 B 细胞受损；将患病鼠的 T 细胞过继转移给健康幼鼠，则可引起类似人类 IDDM 的病理性改变；在患病鼠的胸腺内注射胰岛 B 细胞，可减缓病理损伤。细胞因子的转基因研究近年来取得了很大进展。例如，在胰岛素基因起始序列和终止序列内，替代性的插入 IFN-γ 基因所获得的转基因鼠，其胰岛 B 细胞不仅可分泌 IFN-γ，并且使胰岛 B 细胞高表达 *HLA*-Ⅱ类分子。转基因鼠出现与人类 IDDM 和自发性糖尿病小鼠相似的病理性改变。

（七）自身免疫病的治疗

自身免疫病的治疗目的是抑制异常的免疫应答和恢复免疫耐受状态。例如，对 IDDM 病人可补充胰岛素以代偿胰岛 B 细胞的功能；由于大多数自身免疫病伴有慢性炎症，抑制炎症反应也是治疗的一个重要方面。另外，血浆置换可去除血清中的自身抗体或抗原抗体复合物。近年来，诱导免疫耐受的口服抗原肽、抑制 HLA 功能的药物（如抗 HLA-DP、HLA-DQ 抗体、HLA 自身抗原竞争性模拟肽等），及针对 T 细胞识别和激活的抗体（抗 CD4、CD8、CD3、CD40L 抗体等），虽然目前大多数仍在临床前试验阶段，但仍显示出令人鼓舞的应用前景。

（燕　秋）

第十七章　应激反应的分子机制

关键知识点

- 在全身应激反应中，大量激素的产生可改变许多细胞的代谢和功能状态，各种急性期反应蛋白对机体起保护性作用，细胞内产生的多种热休克蛋白则对细胞产生保护作用。
- 在没有全身应激的情况下，许多因素可引起细胞应激反应，反应过程包括感受应激原信号、启动相关信号转导途径、改变细胞内效应蛋白的活性、促进应激反应相关基因的快速表达、合成多种对细胞具有保护作用的应激蛋白、保护细胞免受损伤或修复已有的损伤、诱导细胞的凋亡等步骤。重要的细胞应激反应有 DNA 损伤应激、低氧应激、氧化应激等。

关键词： 应激原；应激反应；全身应激反应；激素；急性期反应蛋白；热休克蛋白；细胞应激反应；DNA 损伤应激反应；低氧应激反应；氧化应激反应

机体在各种内外环境因素刺激下出现的非特异性适应性反应，称为应激（stress）或应激反应。应激反应属于正常的生理反应，也是生命的最基本特征之一。应激反应是机体的非特异性适应性保护机制，其主要意义是抗损伤，有助于机体抵抗各种突发的有害事件。应激反应对细胞的影响具有两面性：一方面表现在保护细胞免受外界刺激的伤害，另一方面，过于持续或强烈的细胞应激反应又会对细胞造成损伤，具体表现为细胞内各种代谢紊乱，最后导致细胞死亡或凋亡。本章主要介绍全身应激反应中具有重要作用的一些关键分子和主要细胞应激反应的分子机制。

第一节　概　　述

一、应激原是诱发应激反应的因素

能引起机体或细胞产生应激反应的刺激因素称为应激原（stressor）。有些应激原可刺激产生全身性应激反应，有些应激原则主要刺激产生细胞应激反应。应激原大致可分为五类：①物理因素，如电离辐射、紫外线照射、高温或低温以及某些机械性刺激；②化学因素，如活性氧、致癌剂、蛋白质和 RNA 合成抑制剂、重金属以及多种细胞因子等；③生物因素，如病原体（细菌、病毒、寄生虫等）及其产物（如内毒素等）；④细胞营养物质的缺乏，如缺氧、蛋白质缺乏、维生素缺乏等；⑤心理、社会环境因素，如战争、失业、失学、失恋、丧失亲人等。

每一种应激原，必须达到一定强度才能够激发应激反应。不同的人，对于同等强度应激原的反应可以有明显不同，在某些人可引起强烈应激反应的因素可能对另一些人并不起作用，所以无论是全身应激反应还是细胞应激反应，都有着显著的个体差异。

二、全身应激反应是多种信号分子协调体内多种细胞参与的全身反应

全身应激反应是指机体在受到内外环境的各种伤害因素刺激时发生的、多种激素参与的、非特异性的全身反应。这些反应包括以下丘脑-垂体-肾上腺皮质轴和交感-肾上腺髓质

为主的神经内分泌反应和一系列功能代谢的改变，是一组非特异性的全身适应综合征，是机体以最大可能和最有效的方式适应和对抗环境刺激的一种非特异性防御反应。

多数应激反应在应激原作用消除后可很快平静和恢复自稳。但是，有的应激反应表现为一个动态的连续过程，引起疾病，甚至死亡。例如，脑出血可以引起的消化道溃疡；海湾战争之后出现的“海湾综合征”。

三、一系列高度有序的细胞内事件构成细胞应激反应

细胞在生存过程中也不可避免地受到各种内源或外源致损伤因素的作用。对这些有害刺激，细胞会产生防御性或适应性的反应，这种反应称为细胞应激（cellular stress）。细胞应激反应的主要功能是维持细胞正常的结构和功能，从而保证组织器官的正常生存，在无法修复损伤的情况下，应激反应则导致细胞凋亡。

在全身应激反应中，许多细胞的功能和代谢发生改变，参与全身应激反应。然而，即使不存在诱发全身性应激反应的应激原，细胞所处微环境的变化或细胞内发生的事件也可产生特定的应激原，引起细胞应激反应。

细胞应激反应是由一系列高度有序事件所构成，依次为：①细胞感受应激原信号；②启动细胞内应激反应相关信号转导途径；③改变细胞内各种效应蛋白的活性，尤其是一些转录因子的活性；④活化的转录因子促进应激反应相关基因快速表达，合成多种特异性和非特异性的具有保护作用的应激反应蛋白；⑤应激相关的蛋白分子保护细胞免受损伤或修复已有的损伤，最终抵抗应激原的刺激，细胞继续生存；若无法修复应激原造成的损伤，则激活凋亡途径诱导细胞的凋亡。

第二节　全身应激反应的相关分子

在全身应激反应过程中，机体会产生大量的与应激反应相关的分子，协调体内多种细胞发生相应的功能和代谢改变，对应激原刺激产生相应的保护性反应。除了各种细胞产生的与各自功能和代谢改变相关的分子以外，机体在全身性应激反应中会产生一些具有普遍意义的分子。本节介绍三类主要的分子。

一、激素是全身应激反应时产生的细胞间通讯信号分子

全身应激反应过程是一个涉及全身性细胞间通讯和细胞内信号转导的复杂过程，神经-内分泌系统的多种细胞释放出细胞间通讯的信号分子——激素，而更多的细胞则接收信号，通过细胞内信号转导过程调整细胞的代谢及其他各种功能。

（一）应激引起交感-肾上腺髓质系统反应

应激原刺激可引起交感-肾上腺髓质反应，分泌儿茶酚胺。儿茶酚胺除引起中枢效应（主要是兴奋、警觉及紧张、焦虑等情绪反应）外，还可介导外周效应，包括一系列的代谢和心血管变化，如心功能增强、血液重分布、血糖升高。

（二）应激引起下丘脑-垂体-肾上腺皮质系统反应

应激原刺激还可引起促肾上腺皮质激素释放因子和肾上腺皮质激素释放激素（adrenocorticotrophic hormone，ACTH）释放，从而启动下丘脑-垂体-肾上腺皮质（hypothalamo-pituitary-adrenal axis，HPA）轴应激反应。HPA 轴兴奋释放促肾上腺皮质激素释放因子（corticotropin-releasing factor，CRF），CRF 刺激 ACTH 的分泌增加，进而使

血浆糖皮质激素（glucocorticosteroid，GC）的浓度升高。在应激状态下产生的GC可介导多种效应：①促进蛋白质分解和糖异生，补充肝糖原的储备；抑制外周组织对葡萄糖的利用，从而提高血糖水平，保证重要器官的葡萄糖供应。②提高心血管对儿茶酚胺的敏感性，有利于维持血压。③抑制多种炎性介质和细胞因子的生成、释放和激活。④稳定溶酶体膜和减轻有害因素对细胞的损伤作用。

（三）应激引起其他激素分泌的改变

除了上述激素的分泌变化，在应激时分泌增加的激素还有生长激素、抗利尿激素、β-内啡肽、醛固酮等，分泌减少的激素有促甲状腺素释放激素和促甲状腺素。这些激素分泌的变化均引起相应靶细胞的功能和代谢的改变。

全身性应激反应在应激原作用消除后可很快平静和恢复自稳。主要是由于发生改变的激素很快恢复到正常水平。因此，急性应激反应后，机体能够很快恢复。而慢性应激则有可能对机体造成不良影响，其机制主要与激素分泌异常有关。

二、急性期反应蛋白是在非特异性应激反应中大量产生的防御性蛋白

急性期反应（acute phase response）是指各种疾病，特别是感染、炎症、外伤和免疫性疾病时，机体在短时间内（数小时至数天）产生的以防御为主的非特异性应激反应。在急性期反应中，血浆中有多种蛋白浓度迅速变化，这些蛋白质称为急性期蛋白（acute phase protein，APP）。在急性期反应过程中，血浆中任何蛋白质的浓度只要高于或低于正常浓度的25%，都可被定义为急性期反应蛋白。大部分急性期反应蛋白在急性期反应时血浆浓度增加，但也有少数蛋白在急性期反应时减少，称为负急性期反应蛋白，如白蛋白、前白蛋白、运铁蛋白等。

（一）急性期反应蛋白的产生是全身应激反应的组成部分

急性期反应是以防御为主的非特异性应激反应，这种反应的产生实际上是由局部免疫反应诱导的全身应激反应。当机体受到感染、炎症、外伤，或发生免疫性疾病时，局部免疫反应会激活免疫细胞，由免疫细胞产生的细胞因子除参与局部免疫反应外，还可引起肝代谢和基因表达的改变，合成急性期反应蛋白。IL-6和IL-1β是诱导急性期反应蛋白产生的主要分子。IL-1β和IL-6同时也作用于垂体-肾上腺通路，促进ACTH产生，进而诱导肾上腺皮质激素的增加，结果形成一种负反馈机制，因为肾上腺皮质激素可以抑制上述细胞因子的表达。而另一方面，糖皮质激素也可以协同IL-1β等细胞因子的作用，促进肝细胞合成急性期反应蛋白。

（二）急性期反应蛋白基因的表达因应激反应而发生改变

IL-1β和IL-6通过细胞内信号转导途径激活急性期反应蛋白基因。急性期反应蛋白基因启动子和临近的5′非编码区中包含有许多转录因子结合位点，相关的转录因子对急性期反应蛋白基因的表达有促进作用。肝细胞大量合成急性期反应蛋白并将其释放到血液，使其浓度增加50%或数倍，个别重要的急性期反应蛋白的浓度甚至增加1 000倍以上，如C反应蛋白和血清淀粉样P成分。

负急性期反应蛋白在急性期反应时合成减少，主要是针对蛋白质合成能力的代偿性反应，以保证APP的合成能力。

（三）急性期反应蛋白参与全身应激反应

APP的种类很多，分布广泛。表17-1列出了几种重要急性期反应蛋白的特点。急性期反应蛋白的大量产生是一种起动迅速的机体防御机制。机体对感染、组织损伤等的反应

可大致分为两个时相：一为急性反应时相，APP 浓度的迅速升高为其特征之一；另一为迟缓相或免疫时相，其重要特征为免疫球蛋白的大量生成。两个时相的总和构成了机体对外界刺激的保护性系统。

APPs 在机体的防御反应中对机体的保护性作用主要表现在以下几个方面：①抑制蛋白酶的活性。创伤、感染时体内蛋白酶增多，容易造成组织的过度损伤。APP 中包括许多蛋白酶抑制剂，可减少蛋白酶对组织的过度损伤。②清除异物和坏死组织。C 反应蛋白可与细菌细胞壁结合，起抗体样调理作用，并激活补体经典途径，促进吞噬细胞的功能，抑制血小板的磷脂酶，减少其炎症介质的释放。一些转运蛋白可转运、清除有毒物质。补体成分有助于中性粒细胞、巨噬细胞和血浆蛋白在损伤部位的聚集，从而清除局部的病原微生物等损伤因素及死亡细胞的碎片，使组织得以修复。③抗感染：C 反应蛋白、补体成分的增多可加强机体的抗感染能力。纤维蛋白原形成的纤维蛋白在炎症区组织间隙有利于阻止病原体及其毒性产物的扩散。④抗损伤：凝血蛋白类的增加可增强机体的抗出血能力，促进伤口愈合。血清淀粉样 A 蛋白可促使损伤细胞修复。血浆铜蓝蛋白可活化超氧化物歧化酶（SOD），有利于清除自由基和减少组织损伤。

表 17-1　几种重要的急性期反应蛋白

成　分	分子量	正常值（mg/ml）	功　能
血清淀粉样 A 蛋白	160 000	＜10	清除胆固醇
C-反应蛋白	105 000	＜0.5	激活补体
补体成分 C3	180 000	80～120	趋化作用，肥大细胞脱颗粒
结合珠蛋白	100 000	40～80	抑制组织蛋白酶 B、H、L
纤维蛋白原	340 000	200～450	促血液凝固
α1-抗胰蛋白酶	540 000	200～400	抑制丝氨酸蛋白酶
α1-抗糜蛋白酶 1	68 000	30～60	抑制组织蛋白酶 G
α1-酸性糖蛋白	40 000	55～140	促进成纤维细胞生长
铜蓝蛋白	151 000	15～60	减少自由基产生

三、热休克蛋白是在应激状态下保护细胞的蛋白成分

热休克蛋白（heat shock protein，Hsp）是应激反应状态下在多种细胞内产生、对细胞产生保护作用的蛋白质家族。除热应激外，有许多因素可以诱导热休克蛋白的产生，例如缺氧、缺血、砷化物、乙醇、某些病毒感染、氨基酸类似物、化疗药物、环境毒物等。

（一）多种热休克蛋白组成多个家族

根据分子量大小，将主要的 Hsp 分为 Hsp90、Hsp70、Hsp60 及小 Hsp 等数个家族（表 17-2）。此外，还有分子量大于 90 kDa 的大分子 Hsp。各种生物的 Hsp70 具有进化的高度保守性。

（二）热休克蛋白的基本功能是帮助蛋白质正确折叠

Hsp 的主要生物学功能是帮助蛋白质的折叠、移位、复性及降解。其本身不是蛋白质代谢的底物或产物，但始终伴随着蛋白质代谢的许多重要步骤。Hsp 分子中有一个基质识别序列，可与蛋白质分子的疏水结构区相结合。在正常状态下，从核糖体上新合成的蛋白质多肽链尚未正确折叠时，其疏水基团常暴露在外。这些未完成折叠的蛋白质可通过其疏水基团互相结合、聚集而失去活性。Hsp 与这些新合成的多肽链结合，防止其聚集，并帮助其在折叠酶的作用下完成正确折叠。

表 17-2　Hsp 的分类与功能

主要 Hsp 家族成员	分子量（Da）	细胞内定位	可能的生物学功能
Hsp 110 亚家族	～110 000		
Hsp 110		核仁、细胞质	热耐受，交叉耐受
Hsp105		细胞质	蛋白质折叠
Hsp 90 亚家族	～90 000		
Hsp90α（HSP86）		细胞质	与类固醇激素受体结合，热耐受
Hsp 90β（HSP84）		细胞质	与类固醇激素受体结合，热耐受
Grp94		内质网	分泌蛋白质的折叠
Hsp 70 亚家族	～70 000		
Hsc 70（组成型）		细胞质	蛋白质折叠及移位
Hsp 70（诱导型）		细胞质、核	蛋白质折叠，细胞保护作用
Hsp 75		线粒体	蛋白质折叠及移位
Grp78（Bip）		内质网	新生蛋白质折叠
Hsp 60 亚家族	～60 000		
Hsp 60		线粒体	蛋白质的折叠
TriC		细胞质	蛋白质的折叠
Hsp 40 亚家族	～40 000		
Hsp 47		内质网	胶原合成的质量控制
Hsp 40（hdj-1）		细胞质	蛋白质折叠
小分子 Hsp 亚家族	20 000～30 000		
Hsp 32（HO-1）		细胞质	抗氧化
Hsp 27		细胞质、核	肌动蛋白的动力学变化
αB-晶状体蛋白		细胞质	细胞骨架的稳定
Hsp 10	～10 000	线粒体	为 Hsp60 的辅因子
泛素	～8 000	细胞质、核	蛋白质的非溶酶体降解

表中缩写：Bip，免疫球蛋白重链结合蛋白（immunoglobulin heavy chains binding protein）；Grp，葡萄糖调节蛋白（glucose regulation protein）；hdj，人类 Dnaj 类似物（human Dnaj homologue）；Hsc70，热休克同族蛋白（heat shock cognate）；HO-1，血红素氧化酶-1（heme oxygenase-1）；TriC，TCP-1 环形复合物（Tailless complex polypeptide l ring complex）

在应激状态下，各种应激原导致蛋白质变性，使之成为伸展的或错误折叠的多肽链，其疏水区域可重新暴露在外，因而形成蛋白质聚集物，对细胞造成严重损伤。此时大量产生 Hsp，可防止蛋白质的变性、聚集，并促进已经聚集蛋白质的解聚及复性。如蛋白质结构破坏过于严重，无法再解聚及复性时，Hsp 家族成员泛素将会与其共价结合，再经过蛋白酶体将其降解，以保持细胞的正常功能。

（三）热休克因子和热休克元件在应激时调控 Hsp 基因表达

Hsp 的产生有以下特点：①平时表达水平较低，应激时迅速升高；②Hsp 主要存在于细胞质中，应激时 Hsp 迅速进入细胞核和核仁，应激恢复期，又回到细胞质；③Hsp 可被多种应激原所诱导，其表达在转录和翻译两个水平受调控。

热休克蛋白基因的转录调控是通过热休克因子（heat shock factor，Hsf）与热休克蛋白基因启动子区的热休克元件（heat shock element，HSE）相互结合而实现的。Hsf 有 4 种，Hsf 1 是主要的转录因子，在生理和各种应激刺激的条件下，促进相关热休克蛋白的表达。Hsf 1 通常以单体分子的形式定位于细胞质中，应激刺激后，磷酸化的 Hsf 1 转位入细胞核并形成三聚体形式，结合到热休克元件上，诱导热休克蛋白的表达。

Hsp 的 C 端基质识别序列除了可与肽链的疏水结构区结合外，还可结合热休克因子。在热应激或其他应激原作用下，细胞内的蛋白质变性或新生蛋白质折叠出现大量错误，暴露出其疏水区。Hsp 与受损蛋白质的疏水结构区结合，释放出游离的热休克因子，后者聚

合成三聚体，结合到热休克元件上，诱导表达更多的热休克蛋白，帮助蛋白质修复、折叠，或促进无法修复的蛋白质降解。因此，在应激时 Hsp 的生成增多，使细胞维持正常的生理功能，提高细胞对应激原的耐受性。而在恢复期，过多的 Hsp 又可结合热休克因子的单体，抑制三聚体的形成，从而抑制 Hsp 基因的表达，抑制 Hsp 的进一步合成。

第三节　细胞应激反应的分子机制

全身应激反应会引起许多细胞的反应，主要应对机体受到的应激原刺激。而一些细胞内事件或细胞所处微环境的变化也可作为应激原刺激细胞，诱发细胞应激反应。不同的应激原作用于细胞后，细胞会以特定应激方式、通过不同的机制启动细胞内的信号转导途径，引发细胞的各种反应。细胞应激反应是一个复杂的过程，一种应激原常常同时或顺次激活几条信号转导途径，不同应激原也可激活同一条信号转导途径。

细胞应激反应包括多种类型，本节介绍三种细胞应激反应的分子机制：DNA 损伤应激、低氧应激、氧化应激。

一、细胞内 DNA 损伤诱发细胞应激反应

细胞在多种内外因素作用下可发生 DNA 损伤（见第四章），而细胞通过不同机制对 DNA 进行的损伤修复，就是细胞针对 DNA 损伤的应激反应组成部分。

（一）细胞可针对 DNA 损伤产生应激反应

当细胞内的 DNA 发生损伤时，细胞可以感应到损伤的存在，并启动相应的应答反应，即细胞应激反应。细胞对 DNA 损伤的应激反应涉及到多条平行和/或交叉的信号转导途径。DNA 损伤是起始信号，特定的感应分子可感应到这种信号并将其传递给下游的信号转导分子，最终，效应蛋白分子接收到信号（被激活）并发挥各自功能。信号转导网络将启动几种反应：①激活细胞周期检测点激酶，使细胞周期停滞，从而使带有损伤 DNA 的细胞不能继续增殖；②启动 DNA 损伤修复系统，修复损伤，并增强细胞的存活能力；③调控凋亡相关的各种基因的表达。

DNA 损伤后应激反应可导致不同的结局，经过细胞内一系列的应激反应过程，最终会有三种结果：①DNA 损伤不严重，细胞调动所有的损伤修复机制，最终修复被损伤的 DNA，细胞解除周期的阻滞，继续进行正常的生长和分化；②DNA 的损伤比较严重，不能将损伤的 DNA 完全修复，但可以部分修复。修复后的细胞仍然可以进行复制，但有形成肿瘤的倾向；③ DNA 损伤过于严重，无法修复，应激反应即启动凋亡机制，导致细胞发生凋亡。

（二）PI3K 相关激酶感应 DNA 损伤并启动信号转导

细胞内存在着可以识别 DNA 损伤的蛋白质分子，其中最主要的几种蛋白都属于 PI3K 相关激酶（phosphoinositide 3-kinase related kinase，PIKK）家族。PIKK 是一个信号蛋白的家族，属于 PI3K 超家族，在调控细胞生长、基因表达、基因组监视和修复等多方面发挥重要作用。哺乳动物细胞表达 PIKK 家族的六个成员，其中 5 个（ATM、ATR、mTOR、DNA-PK、hSMG-1）是蛋白质丝氨酸/苏氨酸激酶。在细胞 DNA 损伤应激反应中发挥作用的主要有 ATM（ataxia telangiectasia mutated）、ATR（ATM-and Rad3-related）和 DNA 依赖的蛋白激酶（DNA-PK）。

ATM 和 ATR 都能够被 DNA 损伤激活，其确切机制尚不清楚。DNA 损伤后，许多蛋白质因子聚集到损伤位点，ATM/ATR 通常在损伤后几分钟就会出现在损伤部位，其

他蛋白质包括染色质的结构蛋白、染色体修复和维护蛋白。ATM 和 ATR 可能通过与 DNA 相互作用被激活，也可能与修复复合物的成员相互作用而被激活。

DNA-PK 也可以感受到 DNA 的损伤并被激活。DNA-PK 是以相对较高的水平、组成型地存在于细胞核内，其活性通过细胞周期进行调节。DNA-PK 的激活涉及到它与 DNA 和其他蛋白质的相互作用。高迁移率组分蛋白 1（high mobility group proteins，hSMG-1）和 hSMG-2 也可以激活 DNA-PK。

（三）DNA 损伤信号在细胞内传递

细胞 DNA 损伤的应激反应依赖于不同网络的同时激活。激活的信号通过多条信号转导途径向下游传递，这些信号最终被细胞周期调控系统、DNA 修复系统和细胞凋亡调节因子所接收，从而激活细胞周期检测点、DNA 修复、凋亡等应激相关的反应。

1. DNA 损伤应激信号转导激活细胞周期检测点并导致周期停滞　ATM 和 ATR 介导的信号转导可导致细胞停滞在细胞周期的 G_1、S 或 G_2 期（图 17-1），阻止带有损伤 DNA 的细胞继续增殖，使受损的 DNA 或染色体有时间得以修复。ATM、ATR 和 DNA-PK 被激活后，可磷酸化多种下游效应蛋白，如蛋白激酶 Chk1（cell cycle checkpoint kinase）和蛋白激酶 Chk2。这些效应蛋白介导的细胞应答或使增殖停滞在不同的细胞周期，或导致凋亡。ATR-Chk1 和 ATM-Chk2 途径并不是两条平行、完全独立的途径，在 NBS1（Nijmegen breakage syndrome）蛋白的辅助作用下，ATM 也可以使 Chk1 磷酸化。ATM 和 ATR 可通过 Chk1、Chk2 和 P53 介导细胞周期检测点，DNA-PK 则可激活 P53 介导的凋亡途径。而在以 DNA 损伤修复为主的反应期，ATM 途径和 ATR 途径可通过激活 Mdm2（mouse double minute 2），下调 P53，形成一个负反馈调节环路（图 17-1），此时 DNA-PK-P53 介导的诱导凋亡的效应处于次要的地位。

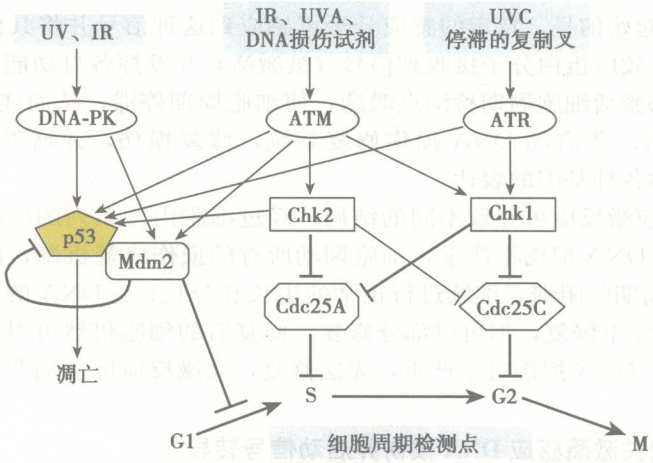


图 17-1　DNA 损伤信号转导途径

ATM 可将 Chk2 磷酸化而激活，Chk2 可将 p53 磷酸化，诱导 G1 检测点。ATR 主要激活 Chk1。ATM-Chk2 途径、ATR-Chk1 途径均可导致 Cdc25A 的降解，激活 S 期检测点。Chk1 激酶调节 DNA 损伤诱导的 G2/M 期停滞。ATR-Chk1 途径是激活 G2 期检测点的主要途径，当 DNA 损伤发生在 G1 或 S 期时，ATR-Chk1 途径可有效发挥作用，使细胞停滞在 G2 期。但如果 DNA 损伤发生在 G2 期，ATM-Chk2 途径将成为主要途径，该途径将 Cdc25C 磷酸化而使细胞停滞在 G2 期。p53 蛋白可通过调控多种因子而导致细胞周期停滞

2. DNA 损伤应激信号转导促进细胞存活并启动 DNA 损伤修复过程　细胞周期阻滞后若 DNA 损伤程度不严重、可以修复时，细胞启动修复系统，修复损伤的 DNA。

（1）DNA-PK 和 ATM 促进细胞存活。DNA-PK 可将 Oct-1 磷酸化而将其激活，启动促进细胞存活的信号途径。ATM 将 NEMO（NF-κB Essential Modulator）蛋白磷酸化，形成 ATM/NEMO 复合物，这种复合物能够激活 NF-κB。在应激反应中，NF-κB 在大多数类型细胞中是作为促生存因子发挥作用。

（2）MAPK 途径促进细胞存活和 DNA 损伤修复。三类主要的 MAPKs（ERK、JNK/SAPK 和 p38MAPK）都可以被各种 DNA 损伤诱导的应激反应信号所激活。如 ATM 可激活 RAF/MEK/ERK 信号途径，亦可通过激活 TAO 激酶（1 种 MAPKKK）而激活 p38MAPK 途径。激活的 MAPK 进入细胞核，将多种靶蛋白磷酸化。MAPK 的靶蛋白包括许多转录因子，激活的转录因子随后调控与 DNA 修复、细胞周期停滞、凋亡相关的多种基因的表达，调节细胞增殖、存活和凋亡。在生长停滞的细胞中，ERK1/2 和 JNK 信号途径均可促进同源重组修复。

（3）多种蛋白质被激活并参与 DNA 损伤修复。ATR 和 DNA-PK 在 DNA 双链断裂位点将组蛋白 H2AX 磷酸化而激活，H2AX 将 DNA 修复蛋白募集到 DNA 损伤位点。ATM-Chk2 途径将 MDC1 蛋白（mediator of DNA damage checkpoint protein）磷酸化而激活，活化的 MDC1 与 H2AX 协同作用。ATM 和 ATR 途径还可将 WRN（Werner syndrome）蛋白激活，WRN 蛋白具有 3→5 核酸外切酶、DNA 解旋酶、DNA-依赖的 ATP 酶活性，参与 DNA 损伤修复。另外还有多种蛋白质与 ATM/ATR 直接或间接相互作用、参与 DNA 损伤修复。

3. DNA 损伤应激信号转导在损伤严重时促进细胞凋亡　当 DNA 损伤严重，修复无望时，细胞启动凋亡信号途径，最终导致细胞的凋亡。Chk1 和/或 Chk2 还可激活 E2F1 和上调 P73，E2F1 和 P73 则转录 *BAX*、*PUMA* 和 *NOXA* 等促凋亡基因。

DNA 损伤应激反应的凋亡信号途径中最重要的蛋白质分子是 P53。P53 蛋白能促进许多凋亡相关基因的表达，如促凋亡基因 *BAX*、*PUMA*、*NOXA*、死亡受体基因 *FAS* 和 *PIDD*、*APAF*-1（apoptotic protein activating factor 1）。P53 蛋白亦可直接从胞质参与凋亡程序，通过促进线粒体释放细胞色素 C，细胞色素 C 与 Apaf-1 蛋白结合，使 Apaf-1 构象发生变化，暴露 N 端胱天蛋白酶（caspase）结合结构域，从而结合并激活 caspase-9，启动细胞凋亡程序。抗凋亡蛋白 Bcl-2 可阻止线粒体细胞色素 c 的释放，而 P53 蛋白则可抑制 Bcl-2 蛋白的表达。

二、低氧可诱发细胞应激反应

缺氧可破坏细胞的有氧呼吸，损害线粒体的氧化磷酸化过程，使 ATP 的产生减少甚至停止，从而引起一系列代谢改变。单纯性的缺氧或缺血可导致低氧性细胞应激，细胞在低氧条件下可感受到氧浓度的降低，产生相应的应激反应。

（一）低氧诱导因子可感应低氧信号而活化

低氧诱导因子-1（hypoxia-inducible factor，Hif-1）是由 α 与 β 亚基组成的二聚体。β 亚基位于细胞核，在细胞中呈组成型表达。Hif-1α 存在于细胞质，受低氧信号调控表达和活化。Hif-1 可调控多种低氧诱导基因的表达，是低氧应激反应中的关键调控分子。

细胞依靠氧感受器感受细胞周围环境的氧浓度变化。NADPH 氧化酶是重要的氧感受器之一。NADPH 氧化酶是一种膜结合蛋白，普遍存在于多种细胞中。NADPH 氧化酶可把细胞周围环境中的 O_2 转变为 H_2O_2。氧浓度高时，H_2O_2 浓度升高，低氧时 H_2O_2 则降低。H_2O_2 的降低有利于还原型谷胱甘肽（GSH）的积累，使许多蛋白质处于还原状态，有助于 Hif-1 的活化。

脯氨酰羟化酶和天冬酰羟化酶也可发挥氧感受器的作用，两者的活性与 O_2 分子浓度成反比。在正常氧浓度时，脯氨酰羟化酶催化 Hif-1α 第 402 和 564 位脯氨酸发生羟基化，这种羟基化的 Hif-1α 经由泛素化而被蛋白酶体降解；天冬酰羟化酶催化 Hif-1α 第 803 位的天冬酰氨发生羟基化，从而阻止 Hif-1α 与 CBP/p300 的相互作用，使得 Hif-1α 无法发挥转录因子活性。在缺氧状态下，上述两种羟化酶的活性降低，Hif-1α 不会被羟基化，降解减少，积累增加，活性增高。

（二）Hif 激活多种基因表达而产生应激反应

当受到缺氧刺激时，在“低氧感受器”的作用下，Hif-1α 亚单位增多，并通过其自身的核定位信号转位入细胞核，在细胞核内与 Hif-1β 亚单位形成二聚体即 Hif-1 分子。Hif-1 是一种重要的转录因子，其靶基因的调控区都有一个被称为低氧反应元件（HRE）的序列。当 Hif-1 结合到 HRE 后，会募集转录辅激活因子 CBP/p300 和转录因子 TIF-2，促进靶基因的表达。

细胞适应低氧环境是通过调控低氧反应基因的表达来实现。低氧反应基因的表达可产生一系列生理适应性反应，如红细胞生成增多，使携氧能力增强；血管增生和血管舒张，增加血液供应；糖酵解能力增强，使无氧条件下 ATP 生成增多，以满足组织细胞的能量代谢，以利细胞和机体在低氧条件下生存。但如果缺氧应激过强、过久则可诱导细胞凋亡。

（三）多种分子参与细胞低氧应激的调控

Hif-1 调控的基因除了与糖酵解和血管生成有关外，还与细胞生长、发育、分化和转移密切相关。当缺氧程度相对较轻时，与糖酵解和血管生成相关的效应分子发挥代偿作用，使细胞维持正常的生命活动。当缺氧程度相当严重时，相关的凋亡效应分子会发挥作用，诱导细胞凋亡。

此外，在低氧信号转导过程中，Ca^{2+}、NO 和 CO 等可能在不同程度上发挥调节作用。低氧时细胞内 Ca^{2+} 升高，激活与钙相关的信号途径，增强低氧敏感基因酪氨酸羟化酶的转录。NO、CO 抑制 Hif-1 的活化，低氧可抑制 NO 合酶及血红素加氧酶的活性，使 NO 和 CO 产生减少，对 Hif-1 的抑制作用降低，促进低氧敏感基因的表达。总之，低氧的信号转导不是一条单一途径，而是多条途径交叉调控的复杂反应系统。

三、活性氧和自由基可诱发细胞应激反应

生物体的氧化应激（oxidative stress）是由活性氧（reactive oxygen species，ROS）或自由基相对超负荷引起的细胞应激反应。体内的氧化应激反应是经常存在的，感染、毒素、代谢紊乱、低温、低氧、局部缺血等因素可造成急性氧化应激，而环境变化、能量需求（如剧烈运动时）、衰老、疾病可造成慢性、长期的氧化应激。

（一）ROS 是诱发氧化应激的主要因素

生物体产生的 ROS 主要指化学性能比氧活泼的含氧化合物，包括超氧阴离子、羟自由基、过氧化氢、过氧化物自由基、一氧化氮和过氧亚硝酸基等。除 H_2O_2 外，以上都是氧自由基，超氧阴离子是产生其他 ROS 的基础。正常情况下，机体利用抗氧化酶如超氧化物歧化酶、过氧化氢酶和谷胱甘肽过氧化物酶等和抗氧化剂如维生素 E、维生素 C、胡萝卜素、谷胱甘肽及微量元素（如硒、锌）等清除体内 ROS 或自由基。若 ROS 或自由基生成增多和（或）清除减少，则可发生氧化应激。

ROS 具有双重作用，低浓度（或适宜浓度）ROS 在细胞内许多信号转导系统中具有

重要的生理作用。当ROS过度产生或抗氧化物不足时，细胞内的促氧化反应和抗氧化反应的平衡被破坏，过量的ROS可损伤细胞内的脂质、蛋白质、DNA，抑制细胞的正常功能。

（二）抗氧化信号转导启动细胞内的抗氧化反应

ROS可直接影响代谢途径中关键酶的活性，通过降低代谢途径的活性可减少ROS的产生。但在ROS过多时，细胞主要通过信号转导途径诱导抗氧化反应。

1. ROS启动信号转导　ROS增多，可引起一些主要信号途径的短暂激活，如MAPK和NF-κB途径。

（1）ROS激活MAPK途径：ROS可激活ERK信号转导途径。ERK5是一种对氧化还原敏感的专一性的MAPK，它的最强激动剂是H_2O_2。凋亡信号调节激酶（ASK1）是JNK和P38分子的上游活化因子，ASK1与氧还原蛋白（Trx）结合而处于无活性状态。ROS可导致ASK1-Trx复合物的解聚，游离的ASK1激活下游的JNK和P38信号传导途径。

（2）ROS通过调节磷酸酶活性而调控抗氧化信号转导：许多丝氨酸/苏氨酸磷酸酶和所有蛋白质酪氨酸磷酸酶的催化结构域中都有一个高度保守序列，其中含有一个对氧化还原状态敏感的关键Cys。这个Cys的氧化可使磷酸酶失活。H_2O_2诱导磷酸酶失活可使NK-κB诱导激酶（NK-κB induced kinase，NIK）等保持磷酸化而处于激活状态。

（3）ROS调节转录因子的合成与降解：ROS可调节一些转录因子的合成与降解，影响其稳定状态时的浓度和/或它们形成同二聚体或异二聚体的能力。例如，转录因子AP-1以氧化还原敏感方式调控许多基因的表达。AP-1的细胞内水平上调和激活都是由MAPK途径介导的。AP-1的一个亚基c-Jun是组成型表达的，需要被JNK磷酸化而激活；另一个亚基c-Fos是诱导表达的，MEK1-ERK途径激活*c-FOS*基因，从而使AP-1的水平升高。

2. 细胞氧化还原状态可以影响信号转导　ROS对转录因子活性的影响是复杂的，它可激活AP-1和NF-κB，但同时也能够氧化这些分子的DNA结合结构域内半胱氨酸残基而降低它们的活性。

3. 抗氧化信号转导途径激活抗氧化基因表达　许多抗氧化基因的启动子区都含有NF-κB和AP-1结合位点，ROS激活的NF-κB、MAPK（活化和上调AP-1）途径最终可诱导氧化应激基因的表达，产生相应的基因表达产物如超氧化物歧化酶、过氧化氢酶和谷胱甘肽过氧化物酶等。谷胱甘肽过氧化物酶和过氧化氢酶可降低H_2O_2的浓度，超氧化物歧化酶上调可加速自由基的清除，对细胞具有特异性的保护作用。

（三）氧化应激反应亦涉及损伤应激信号转导途径

紫外线的照射、电离辐射、化疗药物、细胞因子和毒素等外界因素的刺激也可以诱导ROS水平的升高，诱发氧化应激反应。抗氧化反应只是氧化应激反应的一个部分，如果ROS过多，引起细胞内蛋白质、脂质和DNA的损伤，这些损伤的分子也可以激活细胞内的应激信号转导途径。如DNA损伤后，激活ATM等信号转导途径，产生相应的应激反应。当氧化作用相对较轻时，各种被激活的应激效应分子修复被损伤的DNA、蛋白质、脂质等，恢复细胞的正常生长和代谢过程。当氧化作用过于严重，细胞无法修复时，将启动凋亡机制使受损细胞被清除。

（冯作化）

第十八章　衰老和老年退行性疾病的分子机制

关键知识点

- 衰老及其发生机制：衰老是生物界普遍存在的现象，由遗传决定并受环境影响。现有衰老学说仍不能很好地解释衰老的发生机制。分子生物学认为衰老是分子水平上出现微小变化的综合表现，是基因突变积累的结果。目前科学家们在酵母、线虫、果蝇及小鼠等生物中发现了一些衰老和长寿相关基因。
- 线粒体 DNA 与衰老：目前认为氧自由基随年龄会逐渐增多，累积到一定程度就会损伤 mtDNA，对 mtDNA 的复制、转录及翻译等造成不良影响，导致线粒体结构功能的退化，最终影响 ATP 生成而使机体代谢能力下降，从而发生一系列衰老变化。
- 各种老年退行性疾病及相关基因：危害最严重的是阿尔茨海默病和帕金森病，其共同特征是神经元的退行性病变和凋亡。

关键词： 衰老；阿尔茨海默病；帕金森病；衰老和长寿相关基因；线粒体 DNA

衰老（senescence，senility，ageing）是生物界的普遍现象，指的是生物体达到生殖成熟后在各个水平显现出来的随机的、全身性的紊乱状态。随着衰老的进展，很多慢性疾病的发生也增加，这类疾病称为衰老相关性疾病。尽管衰老是不以人类意志为转移的客观规律，人类从未停止过探求衰老的发生机制以追求更长的生存时间和更高的老年生活质量。本章将主要介绍衰老的概念及其分子机制，同时介绍两种主要的老年相关性疾病，早老性痴呆和帕金森病。

第一节　概　　述

一、衰老是基因突变积累导致的细胞、组织和器官功能衰退

衰老是在生殖成熟后才开始或加速的，进展到一定程度出现明显老年特征就可称之为老年期。关于老年开始的年龄界限自古以来说法不一，西方发达国家长期以来习惯于 65 岁为老年期的开始；1982 年联合国老龄问题世界大会上提出以 60 岁为老年期的开始年龄；我国也一直照此规定。大多数 60 岁以上的人群表现出比较明显的衰老特征。55～60 岁之间患病率最高，慢性疾病也在增加，所以规定老年期始于 60 岁是适宜的，这也同我国民间以 60 年为一甲子的习惯很合拍。随着社会经济和科学事业的发展，人们的健康水平不断提高，有人提出新的年龄划分标准：44 岁以下为青年，45～59 岁为中年，60～74 岁为准老年或称老年前期，75～89 岁为老年，90 岁以上为长寿。

迄今，衰老的本质尚未完全阐明，难以给衰老下个明确的定义，只能就现有认识提出若干概念，从不同角度概括衰老的内涵。分子生物学认为衰老是分子水平上出现微小变化的综合表现，是基因突变积累的结果。

从理论上讲，衰老可分为两类，即生理性衰老（physiological senility）和病理性衰老（pathological senility），前者指机体在其生长的全过程中必然要发生的普通性的退行性变化；后者则主要是由于各种疾病导致的衰老。然而，这两者往往同时存在，很难严格地区

分开来。在大多数老年人身上，这两类衰老互相影响，形成一系列复杂的变化。就每种生物个体而言，衰老速度不尽一致，即有明显的个体差异。一般地说，衰老是一个多环节的生物学过程，早在衰老特征出现之前就已经开始。研究表明，多种因素影响着衰老这一不可逆的发展过程，并且逐步导致机体适应能力降低、储备减少，组织与器官萎缩、重量减轻，最终导致生物体的多种生理功能障碍。

二、分子改变是衰老的基础

解释衰老发生的学说多种多样。古希腊时期的希波克拉底就提出过衰老的温热学说。现代医学和生物学认为除了环境因素的影响，衰老主要由遗传决定，即衰老的发生具有分子生物学基础。以下对衰老的几种学说给予扼要介绍。

1. 程序衰老学说　程序衰老学说（genetic program theory）认为衰老与发育、生长及成熟相似，都是由遗传所决定、按时空顺序表现出来的生命现象。人出生后各器官的生长发育、恒牙对乳牙的替换、胸腺的发育和萎缩、月经的初潮与闭经等也都是按一定程序表现出来的。换言之似乎有个“生物钟”支配着这些生命现象循序展开。实验证明这个“生物钟”在细胞核内。

2. 错误成灾学说　错误成灾学说（error catastrophe theory）认为细胞在复制和基因表达过程中会出现错误，错误的积累和重复导致错误按指数增加，使细胞乃至个体衰老、死亡。DNA 损伤累积所导致的成体干细胞减少是衰老发生的主要机制。成体干细胞对人体自我修复和组织再生至关重要，能分化成机体各种类型的细胞，不断地补充组织细胞，从而实现机体的“自我保持”。DNA 修复酶缺陷的小鼠，其 DNA 修复效率很低，导致造血干细胞 DNA 损伤加快，最终导致干细胞大量消失，小鼠的衰老加速。

3. 自由基学说　自由基学说（free radical theory）意指衰老过程源于自由基对细胞及组织的损害。在生物代谢过程中自由基不断产生，并对自身组织产生毒性作用。

4. 细胞凋亡学说　细胞凋亡与胚胎发育、维持细胞数量恒定、控制细胞增殖、分化及癌变均有密切关系。细胞凋亡受包括 *BCL*-2 和 *P*53 在内的多种基因共同调控，所以也称为程序性死亡。这与程序衰老学说不谋而合，因为程序衰老过程也是由基因编码所控制的生理性主动衰老过程。

5. 端粒学说　人的体细胞的端粒最初的长度大约 17 000 bp，每分裂一次，端粒就缩短 50～200 bp。当端粒缩短到 2 000～4 000 bp 时，正常人双倍体细胞就不能再进行分裂，细胞停滞于 G1 期，开始衰老和死亡。早衰细胞内端粒酶活性低，说明端粒酶的活性与端粒长短有关。将端粒酶活性控制在合适水平，可以延缓衰老，但酶的活性过度激活后使细胞分裂次数增加，成为无限制增殖的癌细胞。如何保持端粒酶的活性且控制在一定临界值之下，有待进一步研究。

第二节　衰老过程受特定的基因控制

衰老过程可能与分化、发育相似，系由早已存在的遗传程序控制，那么细胞中是否存在“长寿基因”、“衰老基因”？近 20 年来，研究者已从多个物种中找到了某些与衰老和长寿相关的特异性基因。除遗传物质外，环境因素对寿命的影响也不容忽视。

衰老基因（ageing gene）和长寿基因（longevity gene）的研究始于简单生物如单细胞的酵母。除酵母外，现已在多种动物如线虫、果蝇甚至在小鼠中发现了一些基因的突变可以延长寿命，这些基因编码的蛋白质常常可以在高级生物中找到类似物；更令人鼓舞的是

科学家在寻找多种动物中均起作用的“公共长寿基因”，这方面已取得一些进展，如在酵母中发现的 *SIR-2* 基因也在线虫甚至哺乳细胞中起调节寿命的作用。与过去认为的衰老发生的多因素学说相比，这些研究提示衰老的发生可能与特异的基因关系更为密切。

一、衰老相关基因影响机体衰老的进程

目前认为，衰老过程可能受特定的基因控制。机体发育成熟之后，基因组内的衰老基因便开始开放，其表达产物则可能直接影响机体衰老的进程。目前，已初步确定某些生物中存在着与衰老相关的特异基因。这些基因大部分是在低等模式生物发现的，其功能尚不清楚，这里只简单叙述部分相关基因。

1. 线虫的衰老相关基因　线虫（c. elegans）是研究寿限与衰老基因的良好模式生物。线虫的寿命一般为 20 天，野生型线虫在 25℃环境中的寿命约为 9 天。线虫的衰老始于发育完成之后，当生殖能力衰退后，线虫的衰老现象日趋严重。已经发现的衰老相关基因包括：①*AGE* 基因，*AGE* 单基因突变可提高平均寿命 65%，最长寿命提高 110%；②*CLK-1* 基因，又称为生物钟基因，其突变株的线虫发育较慢，细胞周期延长，代谢速率减慢，对紫外线的耐受性增强，其寿命可由 8.5 天增至 49 天；③*DAF-2* 基因，*DAF* 基因编码产物与线虫发育相关，*DAF-2* 基因和 *DAF-12* 发生突变时，能使线虫的寿命延长约 5 倍。

2. 哺乳类动物的衰老相关基因　在小鼠发现了一衰老相关基因——*Klotho*（*Kl*）基因，其命名取自希腊神话中掌管生命线之神。该基因的缺陷可在小鼠引起与人在老年易患的动脉硬化、骨质疏松、皮肤萎缩及肺气肿等相同疾病，而且在个体成熟前死亡，表明体内存在与衰老相关的特异基因。人的 *Kl* 基因定位在染色体 13q12 区。

3. 人类的衰老相关基因　人类的衰老相关基因研究难度极大，曾陆续报道过一些相关基因，如衰老协同基因（senescence associated gene，*SAG*）、衰老相关 DNA 合成抑制蛋白（senescent cell-derived inhibitor of DNA synthesis，SDI）和生长停滞蛋白（statin）等，也有过在 1、4、7 和 X 染色体上存在与衰老有关基因的报道，但是上述这些结果目前尚未得到公认。

一些疾病相关基因亦可被视为“衰老基因”。例如，从一种隐性遗传性疾病 Werner 早老综合征定位并分离到的 WRN 基因，位于第 8 号染色体的短臂，编码由 1 432 个氨基酸残基组成 DNA 解旋酶。该病患者多为 20～30 岁青年人，临床上表现为头发花白、皮肤皱缩无弹性等。患者的 DNA 损伤修复和转录均发生异常，细胞传代次数减少，端粒长度迅速缩短，但是该基因突变引起 Werner 综合征的原因目前尚不清楚。1996 年，美国科学家卡伦·肖成功地建立了 β 淀粉样蛋白转基因动物模型，约有半数小鼠脑中形成 β 淀粉样斑块，同时出现学习、记忆能力下降等典型的老年痴呆症状，寿命随之明显缩短。

总之，有关衰老基因的研究仍在不断深入。目前认为，衰老是一个十分复杂的过程，因此衰老不可能由某一个单一基因所决定，而是一连串基因激活和阻抑的综合作用，也是多种基因表达的产物相互作用的结果。在衰老的过程中，某些抗氧化酶类基因和细胞生长调控基因起着举足轻重的作用。

二、生物体存在延长寿命的长寿相关基因

（一）酵母中的长寿相关基因

目前研究发现，影响酵母寿命的基因主要有以下几种：

（1）*LAG1* 基因：*LAG1* 基因是影响酵母寿命的主要基因，如果 *LAG1* 基因发生突变，则酵母的寿命随之缩短。

（2）*RAS* 基因：*RAS* 基因分为 *RAS1* 基因和 *RAS2* 基因。*RAS1* 基因能缩短酵母的寿命，而 *RAS2* 基因则可延长酵母的寿命。*RAS2* 基因编码的 Ras2 蛋白能调节腺苷酸环化酶，增强酵母的营养状态，提高其对环境因素的应激反应能力，降低紫外线照射、热应激和自由基等对酵母的损伤。如 *RAS2* 基因缺失，则酵母细胞的生殖期缩短，抗紫外线照射的能力降低，衰老就会提前到来。

（3）*SIR2*（silent information regulator）基因复合物：引起酵母衰老的一个主要原因是其核糖体 DNA 的重复序列发生同源重组，形成染色体外的核糖体 DNA 环。这些 rDNA 环迅速聚集最终导致酵母细胞的死亡。*SIR2* 基因编码的 Sir2 蛋白是一种 NAD 依赖的组蛋白脱乙酰基酶，能抑制 rDNA 环的形成，进而延长酵母的寿命。同时，有些基因表达产物会对染色体产生毒性作用，Sir2 能使这类基因转录沉默，从而降低核基因组的不稳定性，最终延长寿命。

（4）*HST2*（homolog of SIR2）：*HST2* 基因编码的 Hst2 蛋白是 Sir2 蛋白的同系物。与 Sir2 蛋白相似，Hst2 蛋白亦是一种组蛋白去乙酰基酶，与 Sir2 有相同的延长酵母寿命的作用机制。

（二）果蝇

目前尚未找到果蝇单一的长寿基因。然而，通过限制食物摄入量，已经筛选出具有长寿表型的果蝇品系，该品系的平均寿命和最高寿命均可提高 30%，死亡率明显低于一般果蝇。这个长寿果蝇品系需要超氧化物歧化酶和过氧化氢酶两个基因循一定比例在果蝇体内表达，如果转基因果蝇体内仅分别表达 *SOD* 或 *CAT* 基因，则果蝇的寿命并未见升高。综合来看，氧自由基似有加速衰老的作用，消除氧自由基的酶类似有一定的延缓衰老作用。

（三）哺乳类动物

MHC 基因的 H-2 位点与小鼠寿命有关。小鼠免疫状态与寿命有关，CD8 阳性 T 细胞水平高者，寿命较短，这可能与免疫抑制状态的异常有关。

在哺乳动物细胞，*SIR2* 基因表达 Sir2 蛋白类似物——SIRT1（sirtuin type 1）蛋白，SIRT1 蛋白可以与 P53 结合并使之脱乙酰化，下调其活性，减弱其促凋亡的作用，过表达 SIRT1 可拮抗 DNA 损伤诱导的凋亡。在动物体内提高 SIRT1 的含量，可通过减弱慢性凋亡而延长器官的功能，但也可能有促癌作用。

P66^{SHC} 基因编码 P66^{shc} 蛋白，P66^{shc} 系细胞质中酪氨酸激酶的底物，参与细胞信号从酪氨酸激酶到 Ras 蛋白的传递。P66^{shc} 可提高细胞内活性氧的水平，加快 DNA 的氧化损害，为 P53 依赖性细胞凋亡代谢过程所必需。*P66*^{SHC} 基因纯合缺失小鼠的全身和组织的氧化应激能力降低、血管细胞的凋亡减少，同时和野生型相比，寿命显著延长。

近几年来，有关人类的长寿基因也作了许多探索性工作，但至今尚未见人类“长寿基因”的肯定报告。

第三节　线粒体 DNA 与衰老

有关线粒体 DNA（mtDNA）的结构及其遗传信息传递的详细机制见第二章和第三章，部分线粒体遗传病的发生机制见第十章，本节主要阐述 mtDNA 结构变化与衰老的特殊关系。mtDNA 异常与老年性退行性疾病的关系将在本章第四节讨论。

一、线粒体 DNA 异常是线粒体病的分子基础

细胞内各种活动所需的能量由 ATP 提供，ATP 主要来源于线粒体的氧化磷酸化过程。线粒体功能的异常将影响整个细胞的正常功能从而导致病变。有人将这一类疾病称为线粒体病（mitochondrial disease）。由于大脑、肌（包括心肌）和眼等组织器官对能量的需求较高，因而对线粒体异常最为敏感，成为线粒体病常见的受累组织。线粒体病的研究已从线粒体的形态功能变化进入到分子水平，并提出线粒体 DNA 病（mitochondrial DNA disease）的概念，即线粒体 DNA 结构的异常所导致的疾病或相关疾病。

二、线粒体 DNA 突变是衰老自由基学说的核心

目前多数学者认为随着年龄的增加，线粒体氧化磷酸化系统中产生的氧自由基逐渐增多，累积到一定程度就会损伤 mtDNA，对 mtDNA 的复制、线粒体 tRNA 和 rRNA 的生成以及蛋白质的合成等一系列过程造成不良影响，从而导致线粒体结构的改变与功能的退化，最终影响 ATP 生成，机体代谢能力下降从而发生一系列衰老变化。

1. 衰老时氧自由基生成增加及其对 mtDNA 的损伤　由于 mtDNA 是裸露的，无组蛋白保护，复制、转录速率高，催化 mtDNA 复制的 DNA 聚合酶 γ 不具有校读作用，致使错误率高；与核 DNA 相比 mtDNA 缺乏修复机制，因此 mtDNA 易受自由基等毒性物质攻击而损伤，损伤后又不易被修复而积累。mtDNA 氧化损伤后可造成碱基片段缺失、点突变及插入突变等，以缺失多见，其突变率是核染色体 DNA 的 10～100 倍。

从不同年龄组尸检后的心肌样品中提取线粒体，以 8-OH-dG（8-羟基脱氧鸟苷，一种 mtDNA 脱氧鸟苷的羟自由基加合物）代表自由基的数量作为指标，发现随着年龄增加，8-OH-dG 每上升 1.5%，mtDNA 缺失数上升 1.7%，二者呈正相关，说明自由基随增龄积累，相应 mtDNA 缺失也随年龄增加而累积。

2. 衰老时 mtDNA 的突变　mtDNA 缺失突变以 4 977 bp 和 7 436 bp 的缺失最常见。这两种缺失伴随年龄而增加，它们只聚集在老化的有丝分裂后的细胞内，其中肌组织最易受氧化损害，脂质过氧化程度较其他组织严重，mtDNA 缺失也最多。对人骨骼肌细胞 mtDNA的 4 977 bp 片段缺失情况进行定量研究发现，21 岁时缺失率小于 1/100 000，56 岁时达 1/10 000，78 岁时增至 1/5 000。在衰老细胞中这种缺失的绝对比例可高达 0.1%。

mtDNA 点突变与衰老关系的相关报道少于大片段缺失，某些点突变的累积也可能成为人类衰老的重要因素。例如 mtDNA 第 3 243 位 A→G 突变，成人组织突变率为 0.05%～0.1%，较婴儿组织的 0.01%～0.025%高出 5～10 倍。mtDNA 调控区的 D-loop 和相邻的转录启动序列存在着衰老相关的特异性点突变，如 285T→C，368A→G，383I（插入）和 414T→G。在 14 例年龄超过 65 岁的老人中发现 10 例均有上述点突变，而在 13 例青年中未发现一例具有上述点突变。

第四节　老年退行性疾病的分子机制

大部分的人类神经退行性疾病（neurodegenerative disease）发生在老年期，又称为老年退行性疾病。虽然诱发这些疾病的病因和病变部位不尽相同，但它们都有一个共同的特征，即发生神经元的退行性病变和凋亡，并最终导致个体死亡。该类疾病主要包括阿尔茨海默病（Alzheimer's disease，AD）、帕金森病（Parkinson's disease，PD）、亨廷顿舞蹈

病（Huntington disease）、不同类型的脊髓小脑共济失调（spinal cerebellar ataxias）、齿状核红核苍白球丘脑下核萎缩（dentatorubropallidoluy-sianatrophy）、肌萎缩侧索硬化症（amyotrophic lateral sclerosis）及脊髓肌萎缩症（spinal muscular atrophy）等。

本节介绍两种最重要的老年退行性疾病，即 AD 和 PD 的分子机制。AD 及 PD 主要发生于中老年，随着人口老龄化，患病人数日益增多。目前，仅美国就有 400 万人患有 AD，每年因 AD 死亡的人数近 10 万，医疗费用高达 600 亿美元。我国有关 AD 的流行病学研究尚不完善，一般认为 65 岁以上人群中痴呆的患病率约为 4%。PD 的患病率仅次于 AD，主要发生于中年以上人群，65 岁以上人群中患病率为 2%。

一、早老性痴呆

AD 又称早老性痴呆，是一种常见的老年开始发病的神经退行性疾病。AD 的主要临床症状是记忆缺失（amnesia）、失语（aphasia）和失认（agnosia）；即所谓的三“A”。病理特征是：神经细胞内出现神经纤维缠结（neurofibrillary tangle）、细胞外存在神经（炎）斑（neuritic plaques）或称老年斑（senile plaques）、脑皮层神经细胞减少以及皮层和脑实质动脉的血管淀粉样变性。根据发病年龄和家族聚集性，AD 可分为早发性 AD、晚发性家族性 AD 和散发性 AD。AD 病人往往在 1～2 年内丧失独立生活能力，10 年左右常因感染等并发症而死亡，在西方国家已成为老年人的第四位死因。

自 1907 年 A. A1zheimer 第一次描述 AD 以来，人们一直在探索该病的病因。虽然 AD 的确切病因还不清楚，但可以肯定它与机体的衰老以及遗传有关。在过去的 10 多年中，人们在努力从分子水平寻找 AD 的特征，与 AD 相关基因的不断发现为揭示该病的病因带来了曙光。

（一）位于染色体的 AD 相关基因

目前已确定的位于染色体的 AD 相关基因有 4 个，分别是位于第 21 号染色体上的 *APP* 基因、第 14 号染色体上的早老素-1 基因、第 1 号染色体上的早老素-2 基因和第 19 号染色体上的 *APOE* 基因。

1. *APP* 基因　AD 的重要病理变化是神经细胞，尤其是大脑皮质可见广泛的淀粉样蛋白沉淀形成的老年斑，这种沉淀是一种由 40 个氨基酸残基组成的多聚体肽，称为 Aβ40（amyloid β），对蛋白酶不敏感。沉淀的 Aβ 聚合物对神经元具有毒性作用，可导致神经元的退行性变。

形成这种淀粉样沉淀的肽的前体蛋白是 Aβ 淀粉样前体蛋白（amyloid precursor protein，App）。App 是一种具有膜受体蛋白样结构的跨膜糖蛋白，分子的大部分位于细胞外，羧基端在细胞质。App 合成后，经过 N-和 O-糖基化及酪氨酸硫化，在翻译修饰后或在此过程中于细胞内分泌途径中被裂解。正常情况下，App 由 β 分泌酶裂解为可溶性 Aβ，裂解过程大多发生在 O-糖基化修饰之后。单分子的 Aβ 是一种 β 折叠的可溶性异质的多肽，含有 28～43 个氨基酸残基。

编码 App 的 *APP* 基因在 1987 年被克隆并定位于 21 号染色体长臂 21q11.2～22.1。APP 基因长约 170 kbp，由 19 个外显子构成。目前已发现 *APP* 有多种不同的转录后剪接产物，分别编码 App695、App751、App770、App563 等不同大小的蛋白分子，分子量在 90～140 kDa 之间。其中前三者均含有 Aβ40。Aβ40 在 App695 由第 14、15 外显子产生，在 App770 则由第 16、17 外显子产生。

基因突变或其他因素可以导致 App 氨基酸序列或裂解部位的改变，从而产生易于沉淀的 Aβ。已经发现的 App 突变位点有 Val717Ile、Val717Phe、Val717G1y、Asp670Lys、

Met671Leu 和 Ala692Gly 等。Val717Ile 的突变导致一种长的 Aβ 片段生成；其他几种突变导致 Aβ 生成增加。在 *APP* 转基因鼠，App C 端片段的大量表达，可引起细胞外淀粉样斑块、细胞内神经原纤维缠结和神经元退行性变。

APP 基因突变仅见于少数家族性 AD 患者，如 Val717Ile 见于欧洲和日本的 9 个家族，其他 APP 基因突变仅各见于 1 个家族。通过对家族性、散发性及大量正常人群进行突变基因筛选的结果表明，*APP* 基因的突变仅出现在 2%～3%的早发家族性 AD，所以要阐明 *APP* 基因突变在 AD 发病过程中的普遍作用，还有待深入研究。

2. 早老素-1 基因　早老素-1（*presenilin-1*，*PS-1*）基因定位于第 14 号染色体上（14q24.3），由 10 个外显子构成，由于所研究的家族呈现早年发病，所以该基因被命名为早老素。产物 Ps-1 是一种膜蛋白，也见于 AD 脑内的老年斑内。*PS-1* 基因突变可导致长的、易聚合的 Aβ 分泌增加，其原因可能是影响了 App 的加工、转运及其与细胞骨架蛋白之间的相互作用等，最终导致 AD 的病理变化。

PS-1 基因突变是大多数家族性 AD 的常见原因，占了 55 岁前常染色体显性遗传 AD 病人的 70%。目前已经鉴定的 *PS-1* 基因突变超过 40 种，同时也不能排除在散发性 AD 中有 *PS-1* 基因突变，而且突变也可见于晚发性 AD。

PS-1 基因的突变包括错义突变、剪切位点突变及遗传多态性变化。已经鉴定了 24 种错义突变，均位于 *PS-1* 基因的高度保守序列之中，这些点突变仅存在于早发家族性 AD 患者中。在早发家族性 AD 的家系中还发现了 *PS-1* 基因的错误剪接。*PS-1* 的多态性与晚发 AD 相关，呈非显性遗传模式，*PS-1* 基因 1/1 型与白种人群 AD 的发病率相关，其风险率是 2/2 型和 1/2 型的 2 倍，还具有性别差异，女性高于男性。如果在 *APOE* 基因分型的基础上（见后）对 *PS-1* 再进行基因分型，则非 *APOE*ε4 型患者中，*PS-1* 基因 1/1 型的风险率较 *APOE*ε4 型患者高 1 倍。

3. 早老素-2 基因　早老素-2 基因（*Presenilin-2*，*PS-2*）位于第 1 号染色体 lp31.42，编码由 448 个氨基酸残基组成的蛋白，其结构与 Ps-1 十分相似。在 AD 患者脑中，Ps-2 在基底节、额叶皮层和海马区的表达水平较对照组明显减少。目前 Ps-2 蛋白在正常细胞中的功能尚不清楚，仅知道全长 Ps-2 蛋白可以分解成 2 个片段：一个是 35 kDa 的 N 末端片段，另一个是 20 kDa 的 C 末端片段，后者与细胞骨架调控有关。Ps-2 在细胞内可以通过泛素化途径降解。

PS-2 基因突变发生率低，至今确定有 2 个突变点关联到 AD 发病，仅在 2 个家族性 AD 家系中存在。

家族性 AD 中发现的 *PS-1*、*PS-2* 及 *APP* 基因突变似乎可以解释大多数早发家族性 AD。不过，不具有这些突变的早发 AD 家族的存在仍然提示，至少还有其他早发家族性 AD 的基因尚未发现。

4. *APOE* 基因　*APOE* 基因定位在 19 号染色体 19q13.2 上。全长 3.7 kb，含有 4 个外显子和 3 个内含子。成熟 ApoE 蛋白质的分子量是 42 kDa，由 299 个氨基酸残基组成。人体内 ApoE 的类型主要取决于其三个等位基因编码三个等位体（*APOE*ε2、*APOE*ε3 和 *APOE*ε4）。三个等位体结构类似，仅有个别氨基酸不同：ε3 在 112 位上是半胱氨酸，158 位是精氨酸；ε4 在这两个位点上为精氨酸；而 ε2 则是半胱氨酸。在欧洲，*APOE*ε3 等位基因频率为 0.77，*APOE*ε2 为 0.08，*APOE*ε4 为 0.15。在中国和日本，*APOE*ε4 较少见（频率为 0.06～0.12），而在非洲（0.21）、芬兰（0.23）较常见。

APOE 等位体作用于 Aβ 的形成、聚集和清除。ε4 等位体与 tau 和 MAP2C 蛋白结合强度弱于 ε3；后者与神经原纤维缠结形成有关。ε4 的脂质运输能力差，影响神经细胞丢失后的代偿过程，胆碱能神经递质需要脂质内环境平衡，并影响胆碱乙酰转移酶活性。体

外研究提示 ε3 有促进神经轴突延伸作用，而 ε4 则抑制此作用。ε4 抗氧化酶活性减少，ε2、ε3 抗氧化作用增加。除此之外，ApoE 含量随 *APOE*ε4 等位基因增加而减少。最近发现，携带 *APOE*ε4 的 AD 患者和携带 *APOE*3/3 的 AD 患者与对照相比，ApoE 浓度降低。*apoE* 基因敲除的小鼠，随年龄的增长突触的完整性破坏，突触的可塑性受损。AD 患者海马突触的再生明显受阻也与 ApoEε4、ApoE 水平及可塑性受损相关。

*APOE*ε4 等位基因可能是散发 AD 与晚发家族性 AD 重要的危险因子。通常 *APOE*ε4 纯合子较杂合子的 AD 危险性高，并且发病年龄提前了 3～16 年。大约 65%～75%散发 AD 病例及晚发家族性 AD 病例与 *APOE*ε4 等位基因相关，而早发家族性 AD 与此基因无关。

（二）线粒体 DNA 异常

在 AD 病人中陆续发现的一些 mtDNA 点突变与 AD 病理机制的关系迄今尚不能确认，一种观点认为，mtDNA 点突变可能不是疾病形成的初始因素，但可能与 AD 的首发年龄有关。

AD 病人组的 mtDNA 片段缺失较对照组缺失高 6.5 倍。Marisol 等发现 AD 和衰老在 mtDNA 的缺失突变方面有着不同的特点，即 AD 病人初发病时，mtDNA 缺失水平高，到了 80 岁以后，缺失量下降。造成这种现象的可能原因是，积聚了足够多的 mtDNA 缺失的细胞会死亡，mtDNA 大量缺失的细胞会随细胞的死亡而消失，因而大于 75 岁的 AD 病人的 mtDNA 缺失程度反而会下降。

AD 的发病与许多因素有关，单一因素都不足以解释 AD 的发病机制。越来越多的资料表明，AD 的发病机制可能是环境因素及患者自身因素共同作用的结果。

二、帕金森病

帕金森病（PD）是中老年人的另一种常见的神经变性疾病。其主要病变为脑中黑质、蓝斑、迷走神经背核等具有色素的细胞变性坏死，多巴胺递质生成障碍，造成多巴胺能系统与胆碱能系统的不平衡，而引起一系列临床表现。年轻的 PD 病人遗传倾向大，14%的年轻型 PD 病人有阳性家族史。另外，震颤型 PD 病人的家族中 PD 的发病率更高，而且在许多家系中往往同时存在数个原发性震颤病人，其遗传方式多呈常染色体显性；以强直为主要表现的病人，家族史较少，可能为隐性遗传。

遗传因素在 PD 发病中起着重要的作用。单用环境因素无法解释为什么人群中许多具有相同环境暴露的个体并不都患病，也不能解释为什么 PD 与一些药物的毒副反应有一定的连锁关系。本节将从 PD 相关基因出发，介绍有关分子机制。

（一）位于染色体的 PD 相关基因

1. α-共核蛋白基因　α-共核蛋白（α-synuclein）基因是从一个染色体显性遗传 PD 家系、先经连锁分析定位于第四号染色体的 q21 段，然后克隆获得的，该基因亦被称为 *PARK1*。该基因在此 PD 家系中有 Ala53Thr 突变，但是该基因的突变后来仅在少数 PD 家族中发现。

α-共核蛋白是一种突触前蛋白，其氨基酸序列与非 β-淀粉样蛋白前体（non-beta-amyloid component precursor）相同。后者是 AD 老年斑的成分之一，参与调节淀粉样物质的沉积过程。α-共核蛋白的基因变异可能使其容易发生泛素化，进而沉积而被包裹，形成 Lewy 小体。Lewy 小体中存在高浓度的 α-共核蛋白。在此过程中，脑中的一种泛素连接酶 Parkin 的正常功能被干扰，进而导致细胞凋亡，黑质神经元缺失。

2. *PARKIN* 基因　*PARKIN* 亦称为 *PARK2*，是从日本的 1 个 PD 家系中首先定位和

克隆的。该基因位于第 6 号染色体长臂（6q25.2～27），有 12 个外显子。整个基因长度为 500 kbp。

在常染色体隐性遗传性幼年型帕金森病患者中，发现 *PARKIN* 基因的第 3 到第 7 外显子有不同程度的缺失或点突变。例如，Thr240Arg 和 G1n311 stop 点突变，该突变在部分欧洲的 PD 家系中亦存在。

PARKIN 基因编码的 Parkin 蛋白由 465 个氨基酸组成，主要分布在细胞的高尔基体上。Parkin 属于泛素连接酶，在细胞内参与蛋白质的选择性降解。前述的 α-共核蛋白是 Parkin 作用的底物之一，Parkin 有阻断 α-共核蛋白基因突变所导致毒性的效果。因此，采用 Parkin 基因治疗可能具有治疗 α-共核蛋白基因突变导致的家族性帕金森病的潜在价值。

3. 第 2 号染色体连锁的 PD 基因-*PARK3* 基因　最近，另一家族性常染色体显性遗传，具有 Lewy 体的 PD 的一个基因被定位在染色体 2p13，该位点的外显率为 40%左右，它们多数与德国北部和丹麦有血缘关系。目前该致病基因尚未克隆。

4. *N44A2* 和 *DJ1* 基因　*N44A2* 和 *DJ1* 是最新发现的 PD 易感基因。*N44A2* 基因编码一组核受体超家族，是黑质多巴胺神经元再生和分化的必要因子。在家族性 PD 病人中发现了 *N44A2* 基因两种突变（291T 缺失和 245T→G）；而在散发性 PD 病人（94 例）和年龄相对应的对照组（221 例）中均未检测到这种突变。此种突变导致了 N44A2 的 mRNA水平显著下降且影响酪氨酸羟化酶的转录，因而可产生多巴胺能神经元的功能异常。

DJ1 基因含 8 个外显子，长 24 kb，其主要的转录本编码 189 个氨基酸残基的 Dj-1 蛋白。Dj-1 属于 Thij/PfpI 家族。荷兰一 PD 家系因 *DJ1* 基因外显子 1～5 缺失而不能产生 Dj-1 蛋白；意大利一 PD 家系则发生 L166P 突变。目前 Dj-1 蛋白的功能尚不明确，但有证据显示，Dj-1 蛋白可能作为抗氧化蛋白减缓细胞质中氧化压力，从而保护神经元免受氧化损伤，Dj-1 蛋白的缺乏和突变使其功能丧失或降低。

（二）线粒体 DNA 突变

研究证实，PD 病人脑黑质的线粒体复合物Ⅰ活性降低 37%，而复合物Ⅱ至Ⅳ活性正常。PD 病人脑黑质线粒体复合物Ⅰ多肽中的低分子量亚单位减少，而复合物Ⅱ至Ⅳ却正常。在线粒体复合物Ⅰ所含的 26 个多肽中，其中有 7 个由 mtDNA 编码。对 PD 病人脑（包括黑质）进行的分析并未发现可以引起线粒体复合物Ⅰ活性降低的基因片段缺失或点突变的存在。

迄今未发现与 PD 相关的 mtDNA 特异性突变。尤其是流行病学研究发现帕金森氏病不具有线粒体疾病的母系遗传特点，因此线粒体异常在帕金森病发病中的作用受到怀疑。

（高国全）

第三篇　分子诊断、预防和治疗

以重组DNA为中心内容的分子生物学技术在20世纪80年代初完成了初创阶段，进入了在全世界的普及阶段，并且在多方面不断有新的进展，技术手段日新月异。这些技术为医学研究领域带来了巨大变化，生物技术正在深刻地影响着对疾病的发生和发展机制的研究，对新的诊断、治疗和预防方法的建立以及新的健康理念的发展。生物技术在世界各国已经或正在成为新的产业生长点。

本篇一方面介绍了一些基本的基因操作原理和技术，另一方面较系统地介绍了已经在临床上使用或试用的采用分子生物学技术发展起来的诊断方法、治疗药物、治疗方案以及疫苗等。鉴于对疾病相关基因的认识正日益深入和广泛，设置了对疾病相关基因发现和克隆策略的介绍。最后，为了帮助读者对分子生物学相关药物和方案在临床治疗方面的使用和发展有更好的把握，本部分还专门安排了发现新药物的策略和药物研究流程及相关法规的内容。

第十九章　基因操作

关键知识点

- DNA变性与复性现象是核酸分子杂交实验的依据。DNA变性的本质是双链的解链。DNA变性从开始解链到完全解链，紫外吸收值达到最大值的50%时的温度称为DNA的解链温度（Tm）。热变性的DNA在缓慢降温过程中，两条互补链可重新配对而复性。只要不同的单链分子之间存在着一定程度的碱基配对关系，DNA与DNA及RNA与DNA间就可以有分子杂交。
- PCR技术利用DNA聚合酶在体外复制DNA。其反应体系包括DNA模板、寡核苷酸引物、DNA聚合酶、dNTP和反应缓冲液等；基本反应步骤包括变性、退火和延伸。PCR技术广泛用于获取、分析和改造DNA。RT-PCR还可用于RNA的分析。
- 基因克隆是将基因片段通过无性繁殖转入另一个生物体内的过程，需要对载体、目的基因、DNA连接酶和宿主细胞等进行操作，从而得到一群完全相同的基因片段。基因克隆的基本步骤是用限制性核酸内切酶切割目的基因和载体DNA、用DNA连接酶连接目的基因和载体、将连接产物导入宿主细胞、筛选及克隆化后获得重组DNA。
- 克隆载体是能够携带目的基因在宿主细胞内扩增或表达的DNA分子，一般应具备复制子、单一酶切位点和筛选标志等元件。最常用的载体是质粒DNA。广义的载体既包括能够携带目的基因在宿主细胞内扩增的克隆载体，也包括可用于表达外源基因的表达载体。原核表达载体和真核表达载体有各自的表达调控元件。

- 有多种技术用于DNA和RNA的结构、定性及定量分析。双脱氧法是最广泛使用的DNA一级结构分析方法。印迹杂交技术将DNA或RNA置于固相支持物，再利用标记的特定核酸探针分析待测样品中相应核酸分子的种类和含量。检测DNA的方法称作Southern blot；检测RNA的方法称作Northern blot。PCR技术，尤其是实时PCR技术正广泛用于基因和基因表达产物的定性定量分析。
- 转基因指将外源基因被整合到宿主动物染色体上且能稳定遗传的技术。基因打靶指通过同源重组定向地从细胞染色体上移除（敲除）或移入（敲入）特定基因的过程。

关键词： 解链温度；基因工程；DNA重组；聚合酶链反应；印迹技术；基因打靶；转基因

自20世纪70年代以来，生物技术（biotechnology）的出现和发展正在深刻地影响着疾病发生和发展机制的研究，并使新的诊断、治疗、预防方法不断涌现，也使人们的健康理念发生了改变。生物技术在世界各国已经或正在成为新的产业生长点。

基因工程（genetic engineering）是生物技术最重要的内容之一，它涉及特定基因的制备、分离、鉴定、改造及其在不同生物间的转移等多项与DNA分子相关的技术，因此，通常又将其称为DNA重组（DNA recombination）技术、DNA克隆（DNA cloning）、或分子克隆（molecular cloning），其核心是对DNA分子的操作。为涵盖更为广义的内容，本章采用基因操作（gene manipulation）这一题目，意指所有涉及DNA或RNA操作的技术，或指所有基因工程和基因工程相关技术。本章主要介绍基因操作的原理、用途以及它们的发展过程，一般不涉及具体的实验操作。另外，蛋白质相关分析技术亦属于重要的分子生物学技术，但由于篇幅所限未予以介绍。

本章将首先介绍核酸的理化性质及核酸酶的概念；然后以获取基因作为切入点，叙述经切割、连接、导入、筛选等基因操作步骤而完成的基因克隆或DNA重组全过程。为了更全面地了解相关技术，对如何进行基因结构分析、基因定性定量分析及基因功能研究等亦进行了介绍。

第一节　核酸的理化性质

核酸的结构及成分赋予其一些特殊的理化性质，这些理化性质已被广泛用作基础研究工作及疾病诊断的工具。

一、核酸的一般理化特性

核酸为多元酸，具有较强的酸性。DNA是线性高分子，因此黏度极大，而RNA分子远小于DNA，黏度也小得多。DNA分子在机械力的作用下易发生断裂，为基因组DNA的提取带来一定困难。溶液中的核酸分子在引力场中可以下沉。在超速离心形成的引力场中，不同构象的核酸分子，如环状、线性、开环和超螺旋等不同结构的DNA沉降的速率有很大差异。这是超速离心法可以纯化核酸的理论基础。

嘌呤和嘧啶环中均含有共轭双键，因此，碱基、核苷、核苷酸和核酸在240～290 nm的紫外波段有强烈吸收，最大吸收值在260 nm附近。这一重要的理化性质被广泛用来对核酸、核苷酸、核苷和碱基进行定性定量分析。根据260 nm处的紫外吸收光密度值(optical density，OD)，可以计算出溶液中的DNA或RNA含量。常以OD＝1.0相当于50 μg/ml双链DNA；40 μg/ml单链DNA(或RNA)、20 μg/ml寡核苷酸(oligonucleotide)

为计算标准。根据 260 nm 与 280 nm 的光密度比值(OD_{260}/OD_{280}）还可以判断核酸样品的纯度，DNA 纯品的 OD_{260}/OD_{280} 应为 1.8；而 RNA 纯品的 OD_{260}/OD_{280} 应为 2.0。

二、核酸分子杂交的原理是 DNA 的变性与复性

在某些理化因素（温度、pH、离子强度等）作用下，DNA 双链互补碱基对间氢键断裂，使 DNA 双螺旋结构松散，成为单链的现象即为 DNA 变性。DNA 变性只改变其二级结构，不改变它的核苷酸排列。

在解链过程中，由于更多的共轭双键得以暴露，DNA 在紫外区 260 nm 处的吸光值增加，并与解链程度成一定的比例关系，这种关系称为 DNA 的增色效应（hyperchromic effect），它是监测 DNA 双链是否发生变性的一个最常用的指标。在实验室内最常用的使 DNA 分子变性的方法之一是加热。如果在连续加热 DNA 的过程中以温度对 A_{260}（代表溶液在 260 nm 处的吸光率）值作图，所得的曲线称为解链曲线（图 19-1）。从曲线中可以看出，DNA 的变性从开始解链到完全解链，是在一个相当窄的温度内完成的。在这一范围内，紫外光吸收值达到最大值的 50%时的温度称为 DNA 的解链温度。由于这一现象和结晶体的融解过程类似，又称融解温度（Tm，melting temperature）。在达到 Tm 时，DNA 分子内 50%的双链结构被打开。一种 DNA 分子的 Tm 值的高低与其分子大小及所含碱基中的 G 和 C 所占比例相关。G 和 C 的含量越高，Tm 值越高；分子越长，Tm 越高。Tm 值可以根据 DNA 分子大小及其 GC 含量计算，计算公式为：Tm=69.3+0.41（%G+C），小于 20 bp 的寡核苷酸片段的 Tm 值可用公式 Tm=4(G+C)+2（A+T)计算。

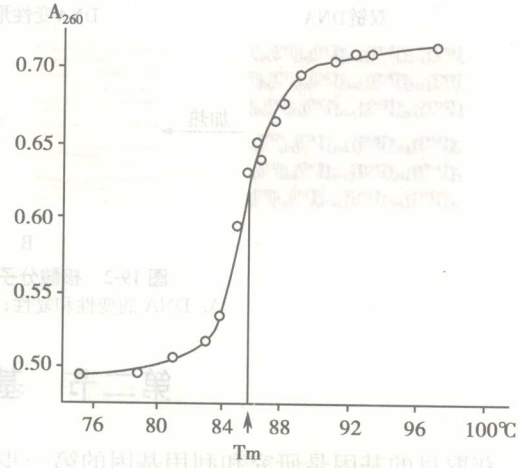


图 19-1 DNA 的解链曲线

变性 DNA 在适当条件下，两条互补链可重新配对，恢复天然的双螺旋构象，这一现象称为复性。热变性的 DNA 经缓慢冷却后即可复性，这一过程称为退火(annealing)。DNA 的复性速度受温度的影响，只有温度缓慢下降才可使其重新配对复性。如加热后，将其迅速冷却至 4℃以下，则几乎不能发生复性。这一特性被用来保持 DNA 的变性状态。一般认为，比 Tm 低 25℃的温度是 DNA 复性的最佳条件。

在 DNA 变性后的复性过程中，如果将不同种类的 DNA 单链分子或 RNA 分子放在同一溶液中，只要两种单链分子之间存在着一定程度的碱基配对关系，在适宜的条件(温度及离子强度）下，就可以在不同的分子间形成杂化双链(heteroduplex)。这种杂化双链可以在不同的 DNA 与 DNA 之间形成，也可以在 DNA 和 RNA 分子间或者 RNA 与 RNA 分子间形成(图 19-2)。这种现象称为核酸分子杂交(molecular hybridization)。这一原理可以用来研究 DNA 分子中某一种基因的位置、鉴定两种核酸分子间的序列相似性、检测某些专一序列在待检样品中存在与否等。分子杂交是核酸研究的重要工具，最新发展出来的基因芯片等现代检测手段的最基本的原理就是核酸分子杂交。

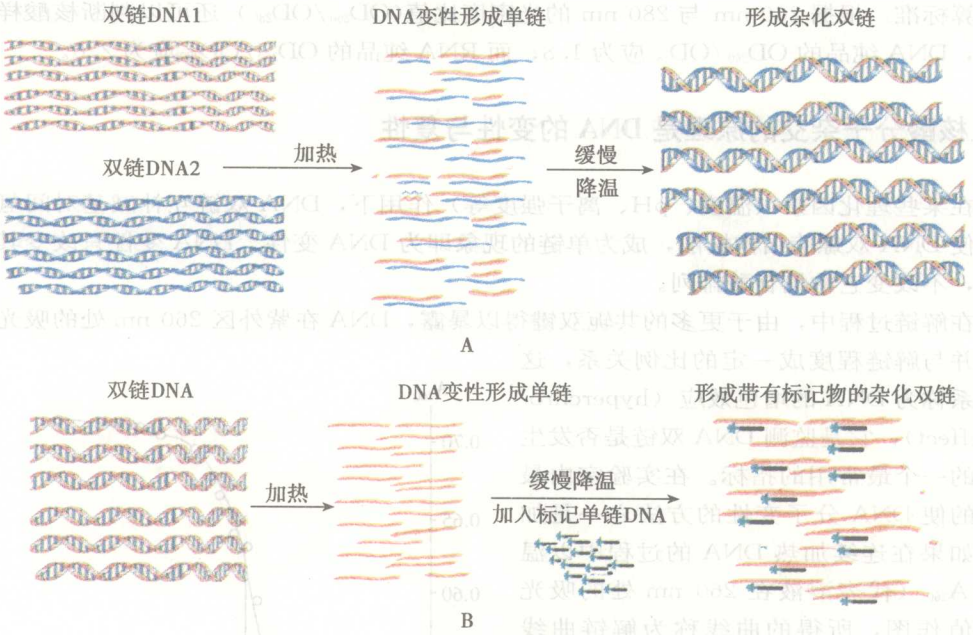


图 19-2　核酸分子杂交原理示意图

A. DNA 的变性和复性；B. 核酸探针用于杂交

第二节　基因的获取

获取目的基因是研究和利用基因的第一步。有多种方法可以获取基因，比如：聚合酶链式反应（polymerase chain reaction，PCR）、反转录-PCR（reverse transcription-PCR，RT-PCR）、文库筛选（library screening）和人工合成等，目前最常用的是 PCR 技术。

一、用 PCR 技术获取基因

PCR 技术是利用 DNA 聚合酶在体外合成 DNA 的方法，其基本步骤包括变性、退火和延伸。该技术根据 DNA 半保留复制原理和 DNA 聚合酶的特性而建立，其基本思路主要来源于：①DNA 的半保留复制机制。②DNA 的变性和复性特征。DNA 的变性可以为复制提供单链模板，而当变性因素去除后，单链 DNA 又能与互补序列重新形成氢键，从而为单链寡核苷酸引物（primer）与模板链的互补结合创造条件。③DNA 聚合酶在单链模板、寡核苷酸引物和核苷酸底物存在下，可以在试管中实现 DNA 合成。与体内 DNA 复制（第三章）不同的是：在 PCR 技术中采用高温使模板 DNA 变性，通过降温使人工合成的寡核苷酸引物与模板互补结合，用耐高温的 DNA 聚合酶（如 *Taq* DNA 聚合酶）催化新链 DNA 的合成，而且通过控制温度的周期性变化，实现 PCR 的循环过程，最后达到特定 DNA 片段扩增的目的（图 19-3）。

虽然 PCR 技术的基本流程很简单，但每一步的基本原理都可能成为实际操作的关键点。为了能更好地掌握 PCR 的技术关键，需要理解以下问题：怎样实现特定长度 DNA 片段的扩增？怎样确定寡核苷酸引物的序列和长度？引物为什么能和模板结合？如何调整 PCR 各步骤的温度和时间？

首先是关于“如何实现特定长度 DNA 片段扩增”的问题。在第三章已述及，DNA 聚

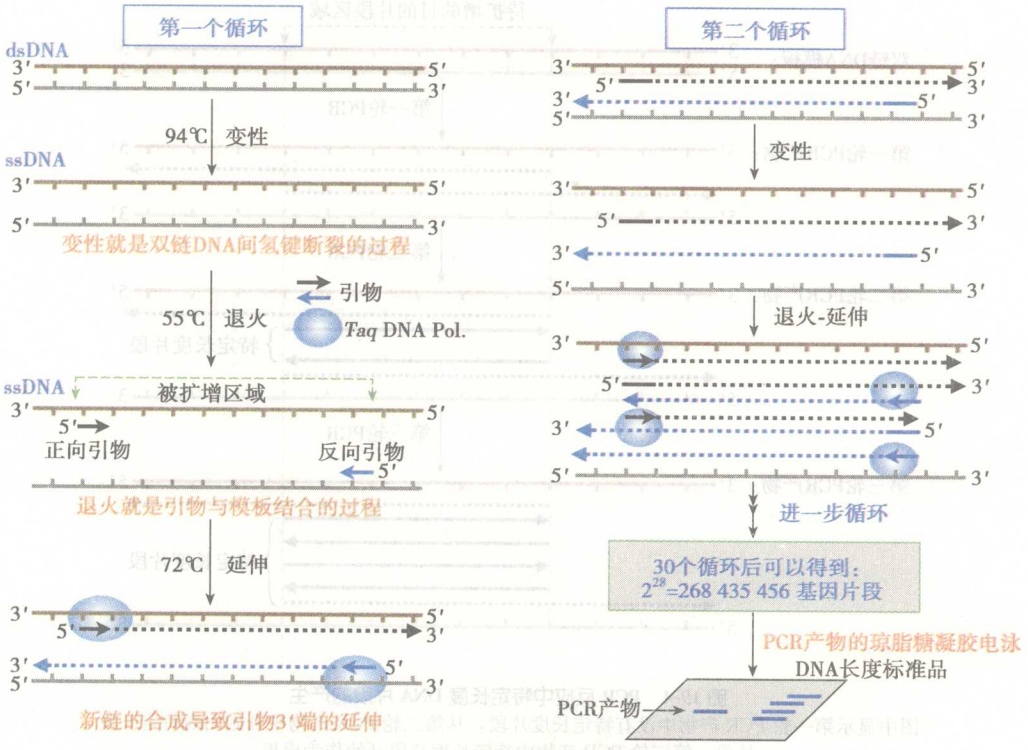


图 19-3 PCR 技术的基本原理示意图

PCR 的三个基本步骤包括变性、退火和延伸。一般在第二个循环后出现特定长度的 DNA 片段，从第三个循环开始，特定 DNA 片段开始作为模板，并逐渐成为扩增反应中的主要片段

合酶催化的 DNA 合成的方向是 5′→3′。如果一对寡核苷酸引物分别与两条互补 DNA 链的特定位置的序列互补，就会在适宜条件下通过形成氢键互补分别结合在两条 DNA 链上。DNA 聚合酶就会以每个引物的 3′-游离端作为添加新核苷酸的位点，从而导致新链以引物的 3′-端为起点、沿 5′→3′方向延伸，因为 DNA 合成的方向不能改变，新合成 DNA 片段的长度就通过引物的位置被确定了，随着 PCR 的不断循环，两条引物之间的 DNA 片段就成了主要产物，也是后续进行 PCR 产物电泳分析可见到的主要条带。应该清楚的是，PCR 扩增产物至少包含 4 种不同长度的 DNA 片段：模板链、2 条引物沿各自模板链合成的两种长片段和以新链为模板合成的特定长度片段（图 19-4）。但是除特定长度片段外的其他产物相对含量很低，对大多数研究没有影响。

引物序列主要依据待扩增 DNA 片段的模板序列设计。一般以正链模板序列确定反向引物(anti-sense primer) 序列，根据互补模板链的序列确定正向引物(sense primer) 序列，从而实现两条引物分别与两条模板链互补结合的目的。一般认为 15～18 个核苷酸就可以保证引物与靶向目标基因结合的特异性。理论上说，引物序列越长特异性应该越高，但错配的发生率也随之上升，因为更长的引物可能在反应体系中形成各种结构造成空间阻碍或在退火过程中分子运动与模板碰撞几率的变化等而影响其与模板互补序列的退火。

采用 PCR 技术获取基因时，一般要求目的基因片段两侧的序列是已知的，这样才能根据已知序列设计引物。引物设计对于获得满意的 PCR 结果非常重要，一般需要考虑引物的长度(至少 15～18 个核苷酸)、与模板的匹配程度、GC 含量（一般为 45%～55%）、适当的末端连接序列、引物自身二级结构特点等，目前已有各种辅助 PCR 引物设计的软件可以帮助将引物设计的更为合理有效，但不能完全依赖软件的设计，最好是人工设计配

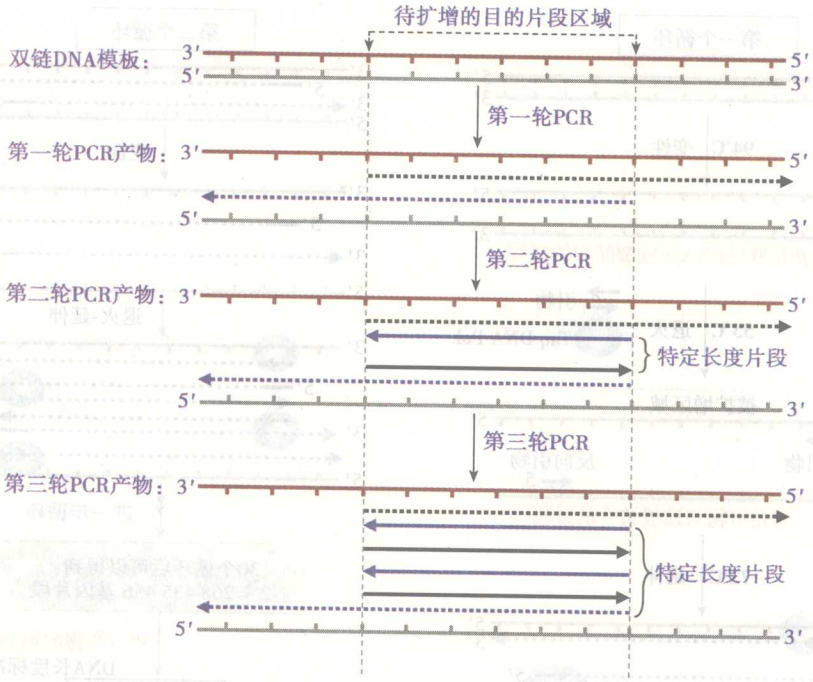


图 19-4　PCR 反应中特定长度 DNA 片段的产生

图中显示第一轮 PCR 产物中没有特定长度片段，从第二轮 PCR 产物中开始出现特定长度片段，第三轮 PCR 产物中特定长度片段开始作为模板

合软件分析相关指数。

第二个问题是引物为什么能与模板结合？基本机制当然是引物序列与模板 DNA 的一部分序列的互补配对。但在 PCR 反应体系中，双链模板 DNA 在变性过程中通过氢键断裂提供单链互补模板链，在退火阶段互补模板可以再次配对结合，引物为什么能找到机会与模板结合呢？实际上，如果模板和引物的数量是一对一的关系，引物确实很难找到机会与模板结合，然而，当引物的数量远远大于模板数量时，情况就会发生变化：在互补模板单链进行分子运动互相靠近的过程中，引物以其分子量小、运动快、数量多的优势抢先与模板上的互补序列发生了结合，结果造成空间位阻妨碍互补模板间的退火，这就是引物与模板退火的基本原理。

最后，PCR 中的各步反应温度是根据什么设定的，是否可以变化调整？PCR 反应过程中的 DNA 变性、引物与模板间退火和新链延伸均需要选择合适的温度，才可高效率地达到扩增 DNA 片段的目的。能完全将氢键断裂但又不影响 DNA 链的磷酸二酯键稳定性的温度是最合适的变性温度，一般认为，90℃以上可以断裂氢键，但最佳效果是在 94±1℃，如果温度超过 100℃，不仅氢键发生断裂，也可影响 DNA 单链的构象，从而影响后续的复性过程。因此，PCR 的变性温度通常在 90～100℃之间（图 19-5）。

退火是引物结合于模板 DNA 的过程，通常采用降低温度为碱基配对创造条件。从图 19-5 中可见，温度越低，配对碱基间越容易形成氢键，一般在 70℃左右时氢键处于断裂和形成的相持状态，因此，退火温度一般应该选择在 70℃以下，55±5℃为实验中常用的经验值，Tm 值减 5℃也是常用的选择。延伸是以引物 3′端为起点、按照模板序列逐个添加核苷酸获得 DNA 新链的过程。为了降低错配的发生，延伸温度一般应选择氢键形成和断裂的相持条件，同时也要考虑 DNA 聚合酶的活性温度，*Taq*DNA 聚合酶的最佳活性温度是 72℃左右，因此，PCR 的延伸温度一般为 72℃。

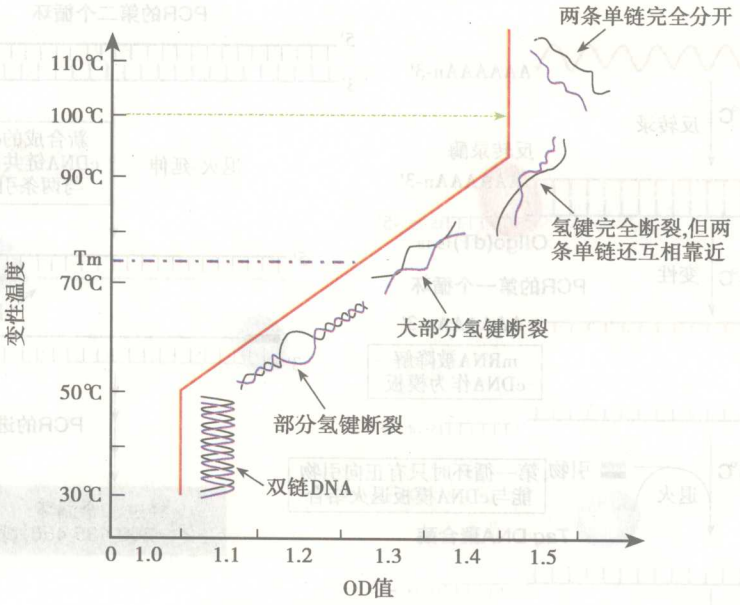


图 19-5　温度对 DNA 结构的影响

DNA 在 50℃以下基本不发生变性；50～70℃时出现氢键断裂现象；一般在 Tm 值附近氢键的断裂和形成处于相持状态，错配的发生几率非常低；90℃以上双链 DNA 间的氢键通常完全断裂

随着生物技术的不断发展，目前以 PCR 技术为基础的其他 PCR 方法也已经越来越成熟，诸如不对称 PCR、巢式 PCR 等，其本质都是利用引物的多种配对方式对片段的扩增，因此，掌握 PCR 的基本原理就可以创建新的 PCR 方法。

二、用反转录-PCR 技术获取基因

反转录-PCR（RT-PCR）是直接以 mRNA 为模板体外合成基因的技术，其基本步骤包括反转录、变性、退火及延伸。RT-PCR 以 mRNA 为模板，先通过反转录过程制备 cDNA，然后再以 cDNA 为模板进行 PCR 合成目的基因（图 19-6）。采用 RT-PCR 可以直接获取基因的编码序列，由于已经去除了内含子区，对于真核生物编码基因的获取比较方便。目前还没有发现耐高温的反转录酶，因此反转录过程通常是在 37～42℃进行，RNA-DNA 杂交链可以直接作为 PCR 的初始模板形式。

三、从基因文库中获取基因

基因文库（gene library）包含基因组文库（genomic library）和 cDNA 文库（cDNA library）两大类。构建基因文库并从中筛选目的基因是获取基因的经典方法，由于基因组序列资源的不断完善，这种方法已经被 PCR 或 RT-PCR 技术所取代，但仍是一种适应特殊需要的一种有效方法，例如需要精确获得数千碱基对的大片段时，PCR 反应过程中存在的碱基错配会造成不便。

基因组文库是指含有某种生物体全部基因片段的重组 DNA 克隆群体。构建基因组文库时，需要先提纯原核或真核细胞的基因组 DNA，然后用机械法或酶法将基因组 DNA 切割成大小不等的许多 DNA 片段，再将其插入到适当的载体（vector）中，继而转入受体菌中进行扩增，从而获得含有多克隆基因组 DNA 片段的混合体（图 19-7）。从理论上说，

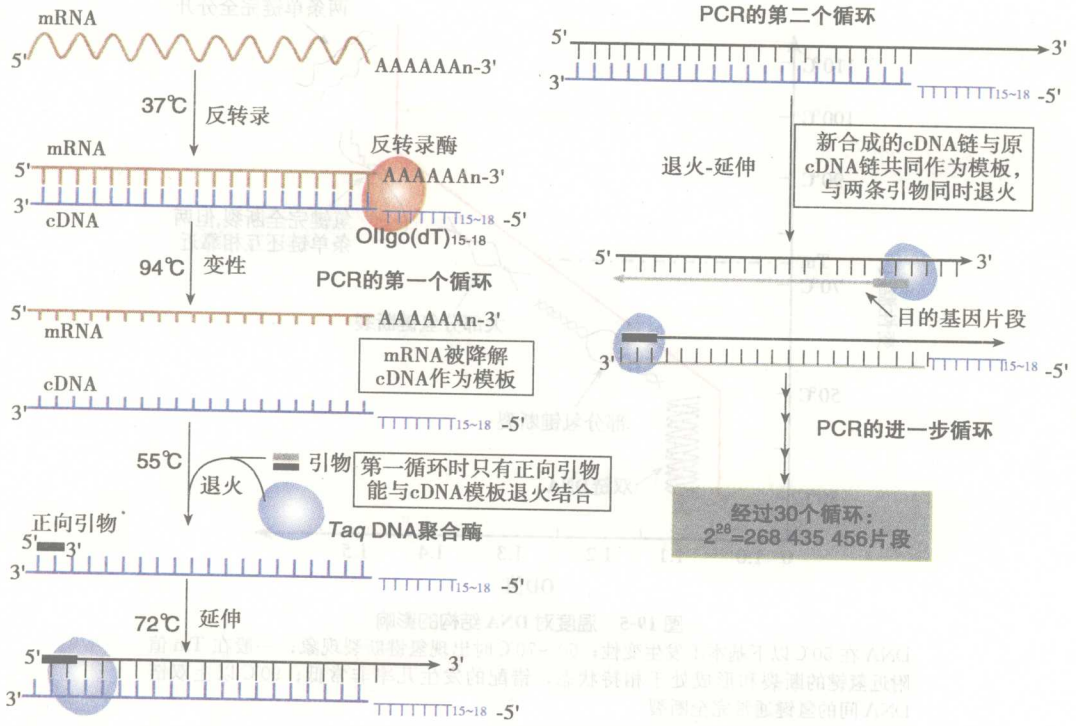


图 19-6　RT-PCR 技术的基本过程

mRNA 作为合成 cDNA 的模板，一旦进入 PCR 流程，mRNA 在高温环境中不稳定而迅速被降解，cDNA 作为 PCR 的起始模板

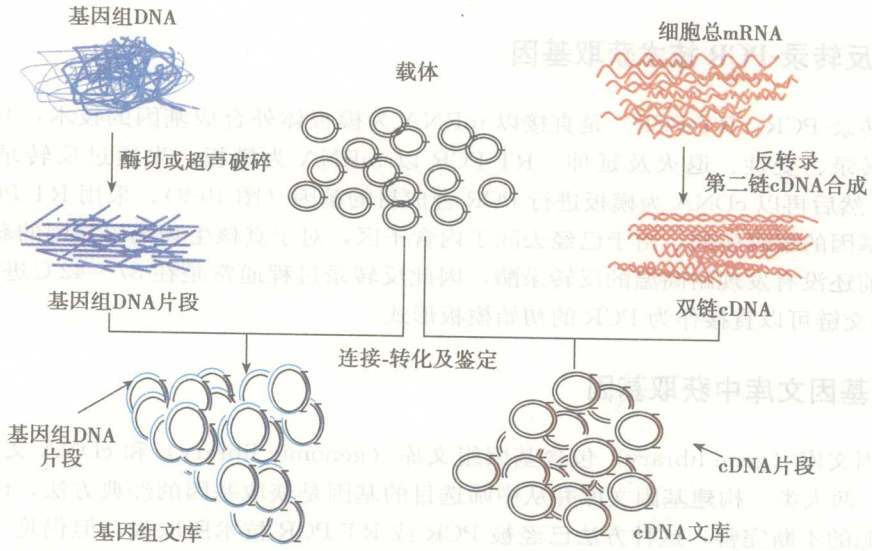


图 19-7　基因组文库或 cDNA 文库的构建

基因组文库可以涵盖基因组的全部基因信息。目前由于各种生物技术公司的服务平台越来越完善，一般是从公司直接购买不同的基因组文库，只有极少情况下才需要自己构建。

cDNA 文库是以细胞总 mRNA 为模板，利用反转录酶合成与 mRNA 互补的 cDNA 单链，再将其复制成双链，然后与合适的载体连接后转入受体菌中所建立的含多克隆 cDNA 片段的混合体（图 19-7）。cDNA 文库理论上包含了一种细胞中的全部 mRNA 信息。

无论是基因组文库，还是 cDNA 文库，在获取基因的操作中仅仅提供了模板，一般可以采用 PCR 技术从文库中获取目的基因片段，也可以采用核酸分子杂交技术从文库中获得目的基因。

四、人工合成基因

人工合成基因有两种方法：一种是直接化学合成基因片段，另一种是化学合成和 PCR 法相结合。对于分子量小的编码序列，可以直接采用化学合成单链 DNA 后体外退火获得双链 DNA；对于较大分子量的基因来说，通常采用分段合成 DNA 片段，然后再利用 PCR 技术将合成的片段经过连接后进行扩增。由于不同生物基因组序列资源的共享，大部分基因序列已经成为已知，因此，利用人工合成法可以很方便地获取目的基因片段，而且利用这种方法可以对基因进行任意的改造和修饰，也可以制造自然界本不存在的基因，或选择合适的密码子等。可见，人工合成基因已经成为一种比较常用的技术。

第三节　基因的克隆

克隆（clone）是指通过无性繁殖过程所生产的与亲代完全相同的子代群体。基因克隆（gene cloning）是指将一个生物体的遗传信息（基因片段）通过无性繁殖转入另一个生物体内的过程，通过基因克隆可以得到一群完全相同的基因片段，由于基因的化学本质是 DNA，因此也称其为 DNA 克隆。通过基因克隆，可以在生物体内将转入的基因片段进行复制扩增。20 世纪 70 年代诞生了 DNA 重组技术，其基本原理是利用 DNA 连接酶能催化两个双股 DNA 分子末端形成 3′,5′-磷酸二酯键的特性，在试管内将目的基因与作为载体的 DNA 分子连接起来形成一个新的嵌合 DNA 分子（即重组 DNA 分子），并在生物活体（或细胞）内获得扩增，也因此使基因重组技术成为分子生物学的核心技术。

基因克隆的必要条件包括：克隆载体、目的基因、DNA 连接酶和宿主细胞。广义的克隆载体（cloning vector）是指能够携带目的基因在宿主细胞内扩增或表达的 DNA 分子。克隆载体一般应具备：①复制子，能在宿主细胞内携带目的基因复制扩增；②一个或多个单一限制性内切酶位点，使目的基因可以插入到载体 DNA 分子中。一般将几个单一限制性酶切位点构建在载体的一个部位，从而构成多克隆位点（multiple cloning site，MCS）；③有筛选标志，即载体 DNA 分子上能赋予宿主细胞一定特性的基因（如抗药基因），用于载体或重组载体的筛选；④用于表达外源基因的表达型载体还应具备能使目的基因转录并翻译的所有调控序列。基因克隆用载体种类繁多（图 19-8），根据用途可以分成克隆载体（cloning vector）和表达载体（expression vector），这里的克隆载体是狭义的，主要是指用于携带目的基因在宿主细胞中扩增的载体；表达载体又可根据宿主细胞分成原核表达载体（prokaryotic expression vector）和真核表达载体（eukaryotic expression vector）；另外，根据载体的来源还可分成质粒载体、病毒载体（包括噬菌体，即细菌的病毒）和人工染色体载体等，目前最常用的载体是细菌的质粒。

基因克隆的基本步骤是：①用限制性内切酶消化目的基因和载体 DNA，使其具备能连接的末端序列特征；②用 DNA 连接酶将目的基因和载体 DNA 连接起来构成重组 DNA 分子；③将连接产物导入合适的宿主细胞（如细菌、酵母、哺乳动物细胞等）使重组 DNA 分子得以复制扩增；④筛选及克隆化，从而获得重组克隆载体及克隆的基因（图 19-9）。

限制性内切酶和 DNA 连接酶是基因操作的重要工具酶。各种Ⅱ型限制性内切酶能专

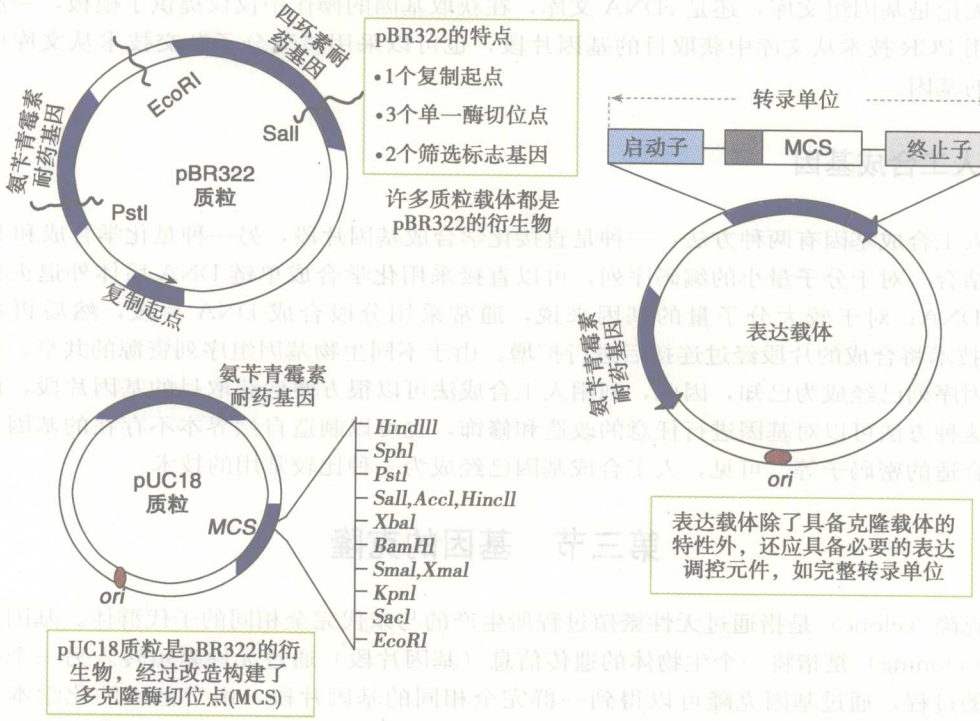


图 19-8　几种克隆用载体举例

pBR322 质粒具备作为载体的基本条件，但没有多克隆酶切位点；pUC18 质粒是在 pBR322 基础上改造的，含有多克隆酶切位点；表达载体具备完整的转录单位

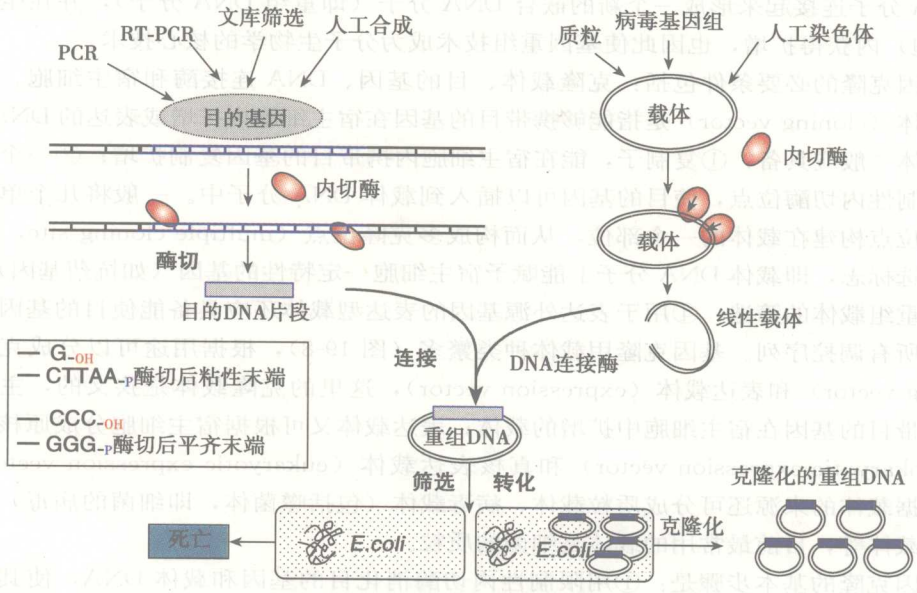


图 19-9　基因克隆的基本流程

一识别特定的 4～6 bp 双链回文结构 DNA 靶序列而成为基因克隆中的常用工具酶。一般情况下，Ⅱ型内切酶在特定位点切割双链 DNA 分子可产生两种不同的末端（图 19-9）：粘性末端（cohesive end）和平齐末端（blunt end）。DNA 连接酶可以将两个双链 DNA 分子连接起来。

通过基因克隆可以对基因序列进行分析，对基因结构进行修饰改造，表达克隆基因的产物，或构建转基因载体或基因剔除载体等。因此，基因克隆是基因操作的核心技术。

第四节　克隆基因的表达

基因克隆的目的之一是获得所克隆基因的表达产物。根据不同用途，可以在原核细胞中表达，也可以在真核细胞中表达。不同的表达系统有各自不同的特点，为了更好地选择合适的表达系统，本节将简介原核表达系统和真核表达系统。

一、原核表达系统

原核表达系统是指由原核表达载体携带目的基因在原核生物中表达目的基因产物，其中最常用的原核生物是大肠杆菌，原核表达载体是大肠杆菌的质粒。原核表达载体除了具备一般载体的特点外，还必须具备原核基因的完整转录单位。原核基因是以操纵子为转录单位的，所以原核表达载体上通常含有操纵子结构，将其中的结构基因用拟表达的目的基因替换即可表达目的基因。影响克隆基因表达的因素主要包括启动子的强弱、RNA 的转录效率、密码子的种类（原核生物偏爱密码子）、表达产物的大小及稳定性等。有些原核生物表达的蛋白质没有生物学活性，可能是由于蛋白质在原核细胞中不能被正确修饰或折叠，遇到这种情况应该采用真核表达系统。

二、真核表达系统

真核表达系统指可使目的基因在真核生物中表达的载体和宿主系统。依据宿主细胞的不同，真核表达系统可分为酵母表达系统、昆虫表达系统和哺乳类动物细胞表达系统等。克隆基因在真核生物中的表达一般需要将基因克隆入真核表达载体中，使目的基因位于真核基因调控元件的控制之下（见第一章）。真核表达载体的各种调控元件大多来源于真核病毒。例如，真核表达载体 pRC/CMV 质粒（图 19-10）的多克隆酶切位点（MCS）上游

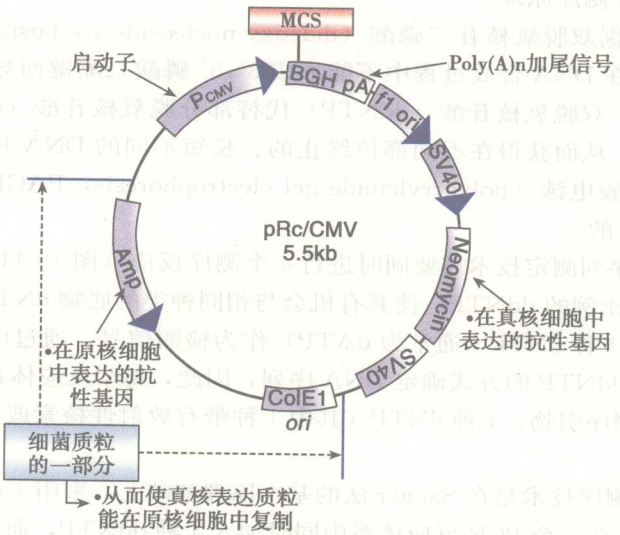


图 19-10　真核生物表达载体结构示意图

的启动子来源于巨细胞病毒（CMV），新霉素抗性基因上下游的调控元件来源于 SV40 病毒。为方便基因克隆操作，现有的真核表达载体又都可以在原核细胞中复制和扩增。这是由于其中有一部分序列来源于细菌的质粒，如原核生物 DNA 复制序列和筛选标志等。这些载体被称为穿梭载体（shuttle vector），既可在原核细胞中复制扩增，又可以在真核细胞中表达目的基因。目前常用的真核表达载体有两类：质粒载体和病毒载体。

利用真核表达系统可以表达目的基因产物，如乙肝病毒表面抗原（HBsAg）是在酵母中表达的；也可以构建转基因细胞模型，将重组真核载体导入真核细胞的过程称为基因转染（gene transfection），受体细胞称作转染细胞。

与原核表达系统相比，真核生物具有翻译后修饰系统，可以对所表达产物进行糖基化等加工修饰，从而获得有功能的活性蛋白。酵母是最简单的真核生物，操作简便，表达产物又可以得到一定程度的加工修饰，是一个用于生物制药的良好系统。

第五节　基因结构的分析及改造

基因的化学本质是一段特定的 DNA 分子，基因的一级结构就是脱氧核苷酸的排列顺序。由于基因的遗传信息蕴含在基因的碱基序列中，因此，如何对 DNA 的一级结构进行分析和改造是本节的主要内容。

一、基因的一级结构分析

基因的一级结构可以采用直接分析法，如 DNA 序列测定法，也可以采用间接法，如酶切电泳分析法，其中 DNA 序列测定技术是分析碱基顺序最直接和最准确的方法。

1. 直接分析基因的一级结构　采用 DNA 序列测定（DNA sequencing）可以直接分析基因的一级结构。目前有两种 DNA 序列测定法，一种是 F. Sanger 发明的双脱氧末端终止法，也称 Sanger 法；另一种是 A. Maxam 和 W. Gilbert 发明的化学裂解法，也称 Maxam-Gilbert 法。两种方法在 PCR 技术建立以后，都在原有技术的基础上引入了 PCR 方法，使序列分析更为快速和敏感，同时为自动化序列分析提供了技术支持。这里以 Sanger 法为例介绍经典的 DNA 测序原理。

Sanger 法是根据双脱氧核苷三磷酸（dideoxy-nucleoside triphosphate，ddNTP）（简称双脱氧核苷酸）在 DNA 合成过程中不能形成 3′,5′-磷酸二酯键而导致链延伸终止的原理，利用 4 种 2′,3′-双脱氧核苷酸（ddNTP）代替部分脱氧核苷酸（dNTP）作为底物进行 DNA 合成反应，从而获得在不同部位终止的、长短不同的 DNA 片段，采用高分辨率变性聚丙烯酰胺凝胶电泳（polyacrylamide gel electrophoresis，PAGE）的分离，实现对 DNA 序列分析的目的。

早期的 DNA 序列测定技术需要同时进行 4 个测序反应（图 19-11），每个反应体系中分别含有 1 种一定比例的 ddNTP，使其有机会与相同种类的底物 dNTP 在 DNA 合成过程中发生竞争，标记 1 种 dNTP（通常为 dATP）作为检测信号，通过电泳、显像的方法阅读 DNA 条带末端 ddNTP 的方式确定 DNA 序列，因此，每个反应体系中的基本成分包括模板 DNA、一条测序引物、4 种 dNTP（其中 1 种带有放射性核素或荧光标记信号）和 1 种 ddNTP。

全自动 DNA 测序技术是在 Sanger 法的基本原理基础上，采用 4 种不同的荧光染料标记 4 种 ddNTP，并在一个 PCR 反应体系中同时加入 4 种 ddNTP，通过荧光检测仪区分不同 DNA 片段末端的荧光信号，经计算机自动分析确定 DNA 序列。目前，DNA 测序工作

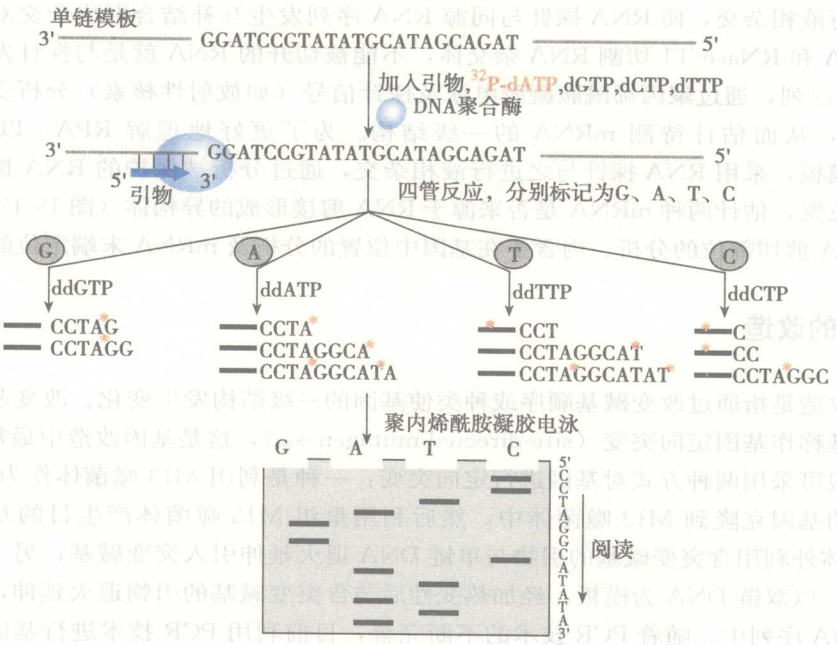


图 19-11　Sanger 法 DNA 序列测定的基本原理

已经成为一种分子生物学的常规服务性项目。

2. 间接分析 DNA 一级结构　采用基因扩增、酶切及电泳等技术，通过比较基因片段的大小可以对基因的一级结构进行间接分析，具体详细方法见基因诊断一章。

3. 在 RNA 水平分析基因的结构　由于基因的直接表达产物是 RNA，因此研究 RNA 的结构可以在某种情况下推测基因的结构，RNA 酶保护试验（RNase protection assay，RPA）就能实现这种目的。RPA 的基本原理是：利用 RNase A 和 RNase T1 能专一性降解单链 RNA 而双链 RNA 受到保护的特性，用体外转录法制备的 RNA 探针与待检测的

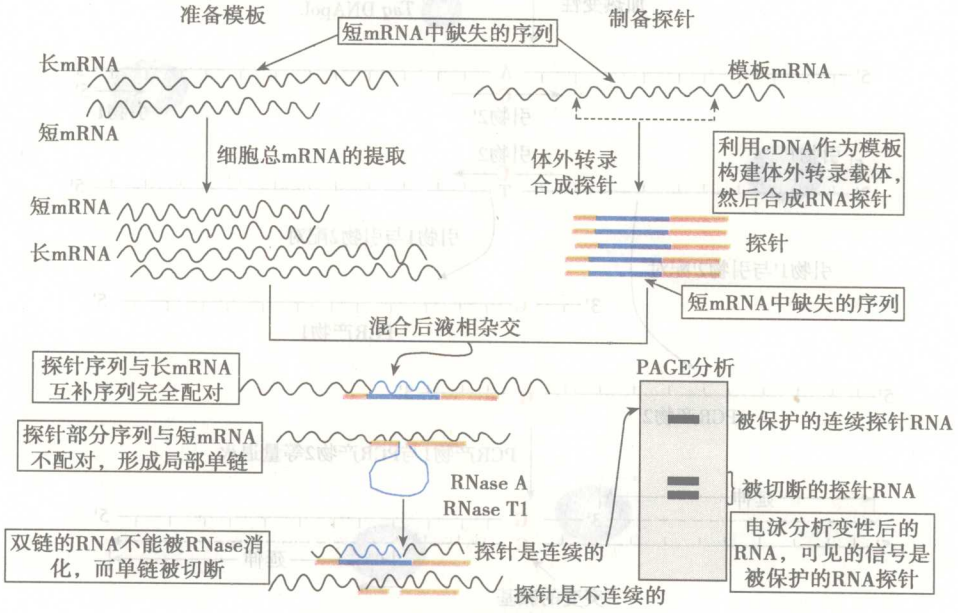


图 19-12　RNA 酶保护试验的基本原理

mRNA进行液相杂交，使 RNA 探针与同源 RNA 序列发生互补结合形成杂交双链，然后用 RNase A 和 RNase T1 切割 RNA 杂交体，不能被切开的 RNA 就是与探针发生互补配对的 RNA 序列，通过聚丙烯酰胺凝胶电泳及探针信号（如放射性核素）分析受到保护的 RNA 片段，从而估计待测 mRNA 的一级结构。为了更好地理解 RPA，以两条同源 mRNA为模板，采用 RNA 探针与之进行液相杂交，通过分析被保护的 RNA 探针在凝胶中的泳动位置，估计两种 mRNA 是否来源于 RNA 剪接形成的异构体（图 19-12）。此法可用于 mRNA 剪切部位的分析、内含子在基因中位置的分析及 mRNA 末端定位的分析等。

二、基因的改造

基因改造是指通过改变碱基顺序或种类使基因的一级结构发生变化。改变基因序列中的特定碱基称作基因定向突变（site-directed mutagenesis），这是基因改造中最常用的技术方法。一般可采用两种方式对基因进行定向突变：一种是利用 M13 噬菌体作为克隆载体，将待突变的基因克隆到 M13 噬菌体中，然后利用重组 M13 噬菌体产生目的基因的单链 DNA，在体外利用含突变碱基的引物与单链 DNA 退火延伸引入突变碱基；另一种是利用 PCR 技术，以双链 DNA 为模板，经加热变性后与含突变碱基的引物退火延伸，将突变碱基引入 DNA 序列中。随着 PCR 技术的不断完善，目前利用 PCR 技术进行基因突变的方法应用更为普遍。

利用 PCR 方法进行基因定向突变的基本程序（图 19-13）是：①设计 4 条引物，其中一对互补引物含有突变的碱基（如引物 2 和引物 2′）；②用两对引物分别进行两组 PCR 反应；③以两组 PCR 产物互为模板和引物，等摩尔比例退火延伸再进行 PCR 反应；④PCR 产物经克隆后测序分析或直接进行测序分析。

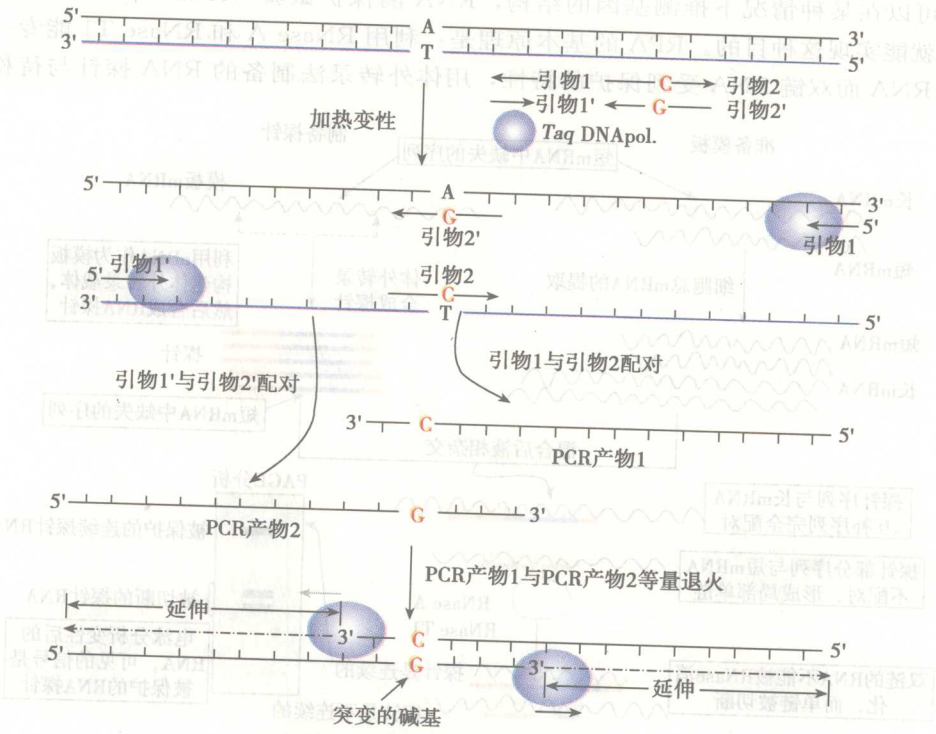


图 19-13　PCR 法定点突变基因流程图

第六节　基因的拷贝数分析

某种基因的种类及拷贝数的分析，即基因的定性和定量分析可以采用 Southern 印迹（Southern blot）技术，也可以采用实时 PCR（real-time PCR）技术，后者可以通过被扩增基因在数量上的差异推测模板基因拷贝数的异同，其基本原理见本章第七节。本节主要介绍 Southern 印迹法。

Southern 印迹的基本原理是：将待测基因组 DNA 经电泳分离后，转移到膜状支持物上使之固相化。利用待测基因中一段已知序列作为探针（probe），依据碱基配对即可复性形成杂交分子的原理，检测固相膜上的待测基因组中与之互补的序列，并根据探针信号出现的位置和次数判断基因的拷贝数。一般情况下，Southern 印迹法可以准确地检测位于基因组不同位置上的相同拷贝基因，但如果基因的多个拷贝成簇地排列在基因组上，应该配合 DNA 序列测定进行分析。

Southern 印迹的基本过程（图 19-14）主要包括：①模板 DNA 的准备。基因组 DNA 需要经限制性内切酶消化为不同长度的片段，电泳分离后转移到固相支持物（如硝酸纤维素膜、尼龙膜等）上。②核酸探针的标记。核酸探针是指核酸分子上带有可被检测的信号分子，如放射性核素或荧光物质等。用于 Southern 印迹的核酸探针包括 DNA 探针、cDNA探针和寡核苷酸探针，极少采用 RNA 探针。核酸探针的标记方法主要有化学法和酶促法。③核酸杂交反应。杂交反应前需变性处理固相支持物上的 DNA 片段（例如碱变性法），使其变成单链 DNA 片段。

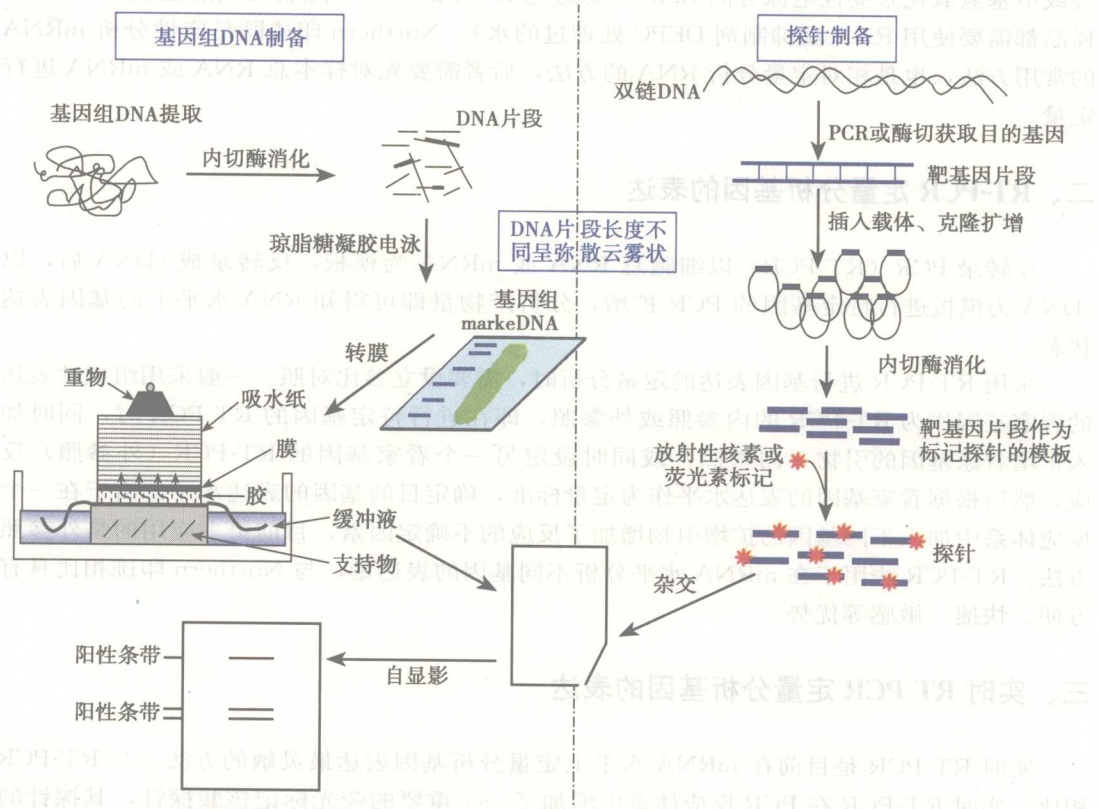


图 19-14　Southern 印迹法实验流程图

Southern 印迹除了作为基因拷贝数的检测方法外，还常用于基因的定位分析、基因酶切图谱分析、基因突变分析和基因重排的分析等，例如，日本学者利根川进利用此法发现了成熟 B 细胞中免疫球蛋白的基因重排现象。

第七节　基因表达的分析

基因表达产物的分析包括在 RNA 水平和蛋白质水平的检测。在 RNA 水平上对基因进行定性或/和定量的分析方法主要包括 Northern 印迹（Northern blot）、RT-PCR、实时 RT-PCR、RNA 酶保护实验和 cDNA 芯片技术等，其中 RT-PCR 是最常用的方法，实时 RT-PCR 是定量分析基因表达的最灵敏和准确的方法。在蛋白质水平上分析基因的表达可以采用 Western 印迹（Western blot）、酶联免疫吸附测定（ELISA）、免疫荧光、流式细胞术、免疫组织化学法等。由于篇幅有限，本节主要介绍在 RNA 水平上定性定量检测基因表达的几种常用方法。

一、Northern 印迹定性分析基因的表达

Northern 印迹是利用核酸探针检测 RNA 的方法，其基本原理和基本程序与 Southern 印迹相似，只是待测分子由 DNA 换成了 RNA。由于 RNA 非常容易降解，RNA 样品的制备和电泳分离就成了 Northern 印迹的关键步骤。在 Northern 印迹中，所采用的探针一般也是 DNA 探针（同 Southern 印迹的探针），极少采用 RNA 探针；一般采用乙二醛、甲醛或甲基氢氧化汞变性电泳分离 RNA，要避免 RNA 的降解（配制反应溶液或溶解 RNA 样品都需要使用 RNA 酶抑制剂 DEPC 处理过的水）。Northern 印迹既是定性分析 mRNA 的常用方法，也是相对定量分析 RNA 的方法，后者需要先对样本总 RNA 或 mRNA 进行定量。

二、RT-PCR 定量分析基因的表达

反转录 PCR（RT-PCR）以细胞总 RNA 或 mRNA 为模板，反转录成 cDNA 后，以 cDNA 为模板进行特定基因的 PCR 扩增，分析产物量即可得知 RNA 水平上的基因表达状态。

采用 RT-PCR 进行基因表达的定量分析时，需要设立参比对照。一般采用组成性表达的看家基因作为 RT-PCR 的内参照或外参照，即在进行特定基因的 RT-PCR 时，同时加入扩增看家基因的引物（内参照）或同时设定另一个看家基因的 RT-PCR（外参照）反应，然后根据看家基因的表达水平作为定量标准，确定目的基因的表达水平。由于在一个反应体系中加入不同基因的扩增引物增加了反应的不确定因素，目前更多应用的是外参照方法。RT-PCR 法用于在 mRNA 水平分析不同基因的表达量，与 Northern 印迹相比具有方便、快捷、敏感等优势。

三、实时 RT-PCR 定量分析基因的表达

实时 RT-PCR 是目前在 mRNA 水平上定量分析基因表达最灵敏的方法。与 RT-PCR 相比，实时 RT-PCR 在 PCR 反应体系中增加了一个重要的荧光标记核酸探针，其探针的 5′端标记了一个报告荧光染料（reporter fluorescence dye），3′端标记了一个猝灭染料

(quencher dye)，两种染料相互靠近（或在同一个探针分子上）时，报告荧光染料不发光，一旦猝灭染料从探针分子上解离下来，报告荧光染料就可以发光。因此，实时 RT-PCR 的基本原理是：利用探针荧光猝灭或发光的特性和 *Taq* DNA 聚合酶外切酶的活性，以 cDNA为模板，在引物引导下，*Taq* DNA 聚合酶开始合成 DNA 新链。新链在模板上延伸时将遇到预先结合到模板上的荧光报告探针，*Taq* DNA 聚合酶就发挥其 5′→3′外切酶活性，将探针从 5′端开始切割，报告荧光染料将从探针上被释放且被专用仪器所监测。在 PCR 的每一轮扩增中释放的报告荧光染料和猝灭染料可以形成一种序列特异性荧光信号，并随着 PCR 的继续而实时增加，所检测的荧光信号也会按比例地增加，通过荧光强度就可以计算出模板 cDNA 的含量（图 19-15）。

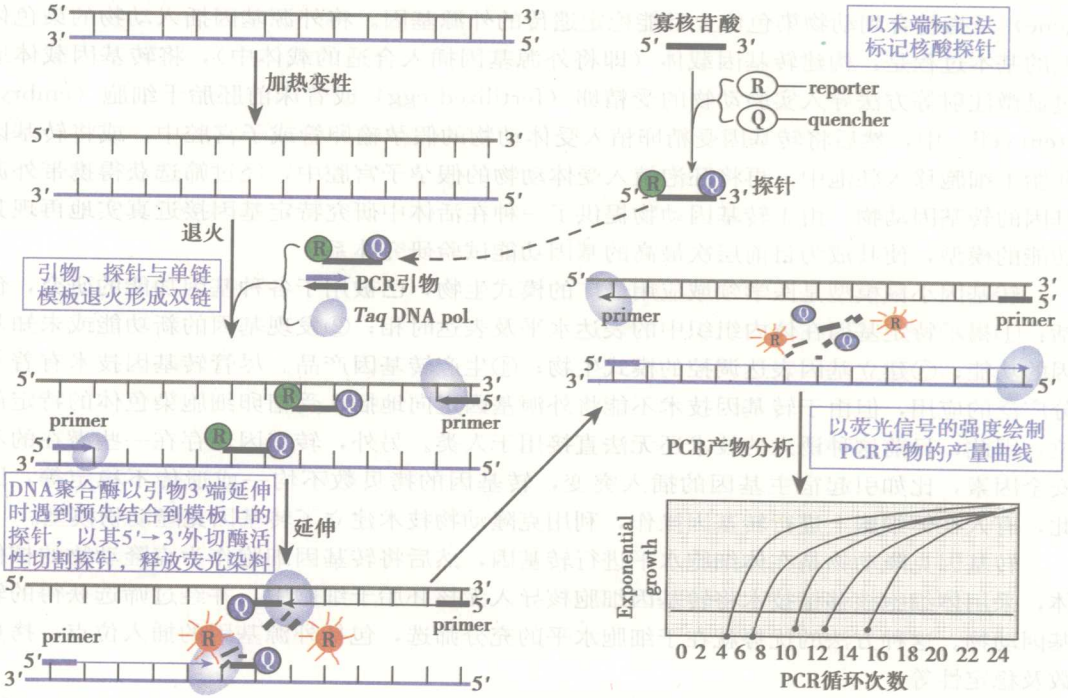


图 19-15　实时 PCR 原理图

实时 RT-PCR 较传统 PCR 有两个明显优势：

（1）实时 RT-PCR 可以在 PCR 进展过程中监测每一个 PCR 循环后的 PCR 产物积累情况，并可对反应产物进行自动化定量分析；而传统 PCR 在反应起始阶段模板浓度相对很低，扩增产量以一个恒定的指数率上升，一旦进入平台期后，扩增效率降低接近零。

（2）实时 RT-PCR 在扩增反应完成后不需要将 PCR 产物从管中取出电泳分析，而是直接通过分析荧光强度判断 PCR 产物量，这样可以减少污染的机会，降低假阳性几率；而传统 PCR 需要将 PCR 产物取出进行凝胶电泳分析。

第八节　基因功能的分析

基因功能的分析可以通过基因表达产物的功能或活性进行推测，但更准确的方法是将基因放到生命体中进行研究，因为基因的功能原本就是在复杂的生命体中实现的，所以，基因功能的研究极大程度上依赖模式生物的发展。转基因技术和基因打靶技术极大地推动

了模式生物的研究进展，各种转基因模型和基因剔除模型为特定基因的功能研究创造了条件。近年来发展起来的 RNA 干扰技术提供了在 RNA 水平上研究基因功能的方法。

一、转基因技术用于基因功能的研究

转基因技术就是将外源基因转移到受体细胞的染色体上，使转基因在受体细胞中表达并发挥其功能的基因操作方法。利用转基因技术可以建立转基因动物模型或基因转染细胞模型。

1. 转基因动物模型用于基因功能的研究　转基因动物（transgenic animal）是指应用转基因技术（transgenic technology）培育出来的、携带转基因的动物。转基因（transgene）是指整合到动物染色体上且能稳定遗传的外源基因。将外源基因插入动物的染色体上的基本过程是：构建转基因载体（即将外源基因插入合适的载体中），将转基因载体通过显微注射等方法导入实验动物的受精卵（fertilized egg）或着床前胚胎干细胞（embryo stem cell）中，然后将转基因受精卵植入受体动物的假孕输卵管或子宫腔中，或将转基因胚胎干细胞移入胚泡中，再将胚泡植入受体动物的假孕子宫腔中，经过筛选获得携带外源基因的转基因动物。由于转基因动物提供了一种在活体中研究特定基因接近真实地再现其功能的模型，使其成为目前层次最高的基因功能试验研究体系。

转基因小鼠模型是医学领域应用最广的模式生物，已被用于各种基因功能的研究，包括：①揭示特定基因在体内组织中的表达水平及表达时相；②发现基因的新功能或未知基因的功能；③建立基因表达调控的模式生物；④生产转基因产品。尽管转基因技术有着十分广泛的应用，但由于转基因技术不能将外源基因定向地插入受精卵细胞染色体的特定部位，因此，目前这种诱人的技术还无法直接用于人类。另外，转基因还存在一些潜在的不安全因素，比如引起宿主基因的插入突变，转基因的拷贝数不均一或遗传不稳定等，因此，有人在体细胞上进行转基因操作，利用克隆动物技术建立了转基因克隆动物模型。

转基因克隆动物是在体细胞水平进行转基因，然后将转基因细胞作为克隆动物的核供体，采用体细胞核移植技术将转基因细胞核导入无核胚胎干细胞中，并经过筛选获得的转基因动物。这种方法的优势就在于细胞水平的充分筛选，包括外源基因的插入位点、拷贝数及稳定性等。

2. 基因转染细胞模型用于基因功能的研究　基因转染细胞（gene transfection cell）是在细胞水平上经过转基因操作获得的细胞模型，其基本流程是：构建外源基因转染载体，然后将其导入细胞中，通过筛选获得表达外源基因的细胞。

利用转染细胞可以直接对基因的功能进行研究，由于目的不同，对转染细胞模型的要求也不同。如果只需要转染细胞短暂地表达目的基因，一般不需要对转染细胞进行筛选，这种转染称作瞬时转染（transient transfection）。瞬时转染时，一般外源基因的表达及其对细胞的影响在 48 小时至 72 小时达高峰，以后由于未接受基因转染细胞的大量扩增，转染细胞很快成为非优势生长而消失。要获得持续表达外源基因的细胞株，必须进行稳定转染（stable transfection），即对转染细胞在压力选择（如加入抗生素）下培养，不含有导入基因的细胞在选择中死亡，获得了基因的细胞生存下来并获得稳定表达外源基因的克隆细胞株。有些基因转染后对细胞有毒性，很难获得稳定转染细胞株。在这种情况下可以选择含诱导型启动子的表达载体，使目的基因在转染细胞内保持关闭状态，只有加入诱导剂后基因才开始表达，这种表达方式称为可诱导型表达（inducible expression）。

二、基因打靶技术用于基因功能的研究

基因打靶（gene targeting）是指通过同源重组定向地从细胞染色体上移除或移入特定基因的过程，一般将定向移除特定基因的过程称作基因敲除（gene knockout），定向移入基因的过程称作基因敲入（gene knock-in）。基因敲除小鼠是目前最常用的基因敲除动物模型。本文主要介绍基因敲除技术的基本原理。

基因敲除是基因体外重组和体内重组相结合的典型例子，利用基因体外重组技术构建基因打靶载体，然后将打靶载体导入受体细胞中使之与染色体上特定基因序列发生同源重组，从而置换特定基因使之活性丧失。

基因打靶载体（图 19-16）包含着待敲除目的基因片段、一个阳性筛选标志基因（如新霉素耐药基因 Neo^R）和一个阴性筛选标志基因（如单纯疱疹病毒胸腺嘧啶激酶的编码基因，*HSV-tk*）。阴性筛选标志基因（如 HSV-tk，也称自杀基因）可以保证非同源重组的细胞在含其底物时被杀死，从而起到负性筛选的目的。

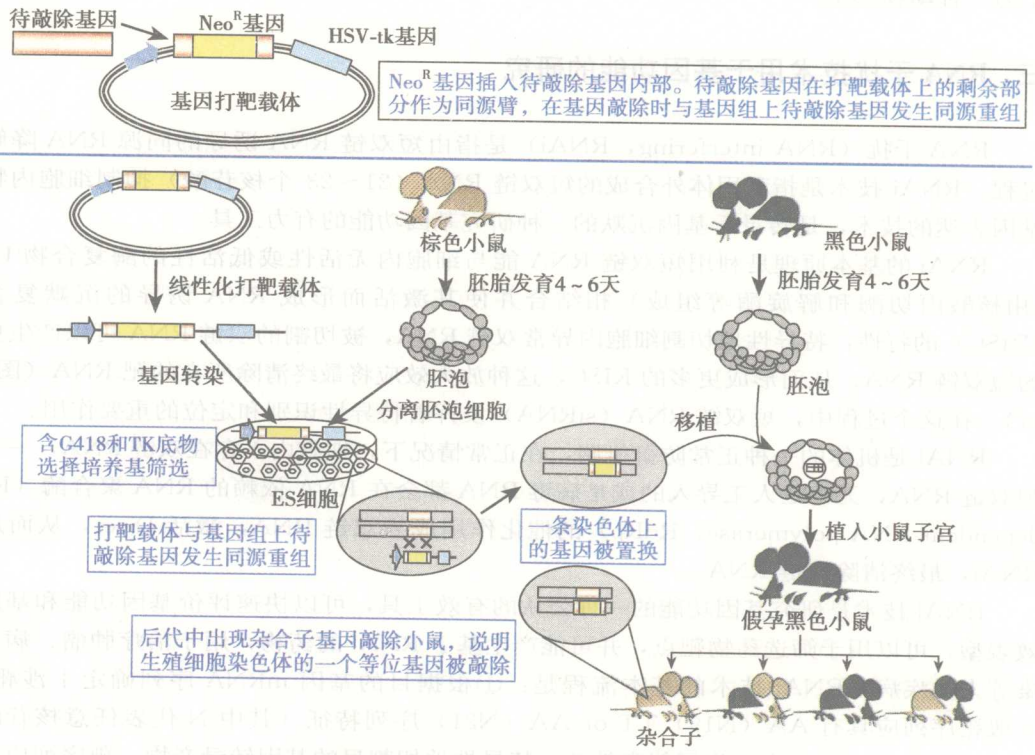


图 19-16 基因打靶用载体的结构及基因敲除小鼠制备原理

打靶载体上待敲除基因的部分序列作为同源臂与基因组上的特定靶基因发生同源重组，结果导致载体上的同源臂序列及其之间的阳性筛选标志基因与基因组上的靶基因发生置换，从而使该基因丧失活性，也同时将阳性筛选标志基因插进了基因组中，这意味着，发生同源重组的细胞具备了阳性筛选标志基因赋予的特性。如果利用胚胎干细胞进行基因敲除，可以进一步通过动物繁育过程获得基因敲除动物。

基因敲除小鼠（图 19-16）是指利用同源重组原理及生物遗传规律被定向地在活体内移除基因的小鼠。需要在体外胚胎干细胞中（embryonic stem cell，ES 细胞）先敲除一条

染色体上的基因，再将发生同源重组的 ES 细胞导入小鼠胚泡中，然后将胚泡植入假孕小鼠子宫腔中。因为同源重组只发生在一条染色体上，获得的是杂合子小鼠，通过小鼠繁育传代后可敲除同源染色体上的等位基因，获纯合子基因敲除小鼠。

随着研究的进展和各种基因调控元件的灵活运用，基因敲除技术也有了新的发展，其中最重要的进展是进行条件性基因敲除（conditional gene knockout）。条件性基因敲除小鼠指专一地在特定组织或特定条件下进行的基因敲除，一方面避免了被敲除基因对胚胎发育过程的影响，另一方面可以直接了解目的基因在不同组织中的功能。实现条件性基因敲除的技术基础是利用了一种重组酶 Cre（Cre recombinase）构成的 Cre/loxP 系统。将该酶的识别序列 loxP 序列（34 bp）插入基因组中待敲除基因的两端，利用组织特异性启动子控制 Cre 重组酶的表达，就可实现组织专一性基因敲除。

如果将 Cre 重组酶基因放在可诱导型启动子的下游，还可通过诱导剂控制 Cre 重组酶的表达时间和表达量，实现基因在不同时间或生长发育不同阶段的定向敲除，根据基因敲除小鼠的各种表型推测被敲除基因的功能。这种方法特别适合于胚胎发育期间一些重要基因功能的研究，也有利于观察基因在不同生长发育期的功能变化，为基因功能的研究提供了另一种动物模型。

三、RNA 干扰技术用于基因功能的研究

RNA 干扰（RNA interfering，RNAi）是指由短双链 RNA 诱导的同源 RNA 降解的过程。RNAi 技术是指利用体外合成的短双链 RNA（21～23 个核苷酸）抑制细胞内特定基因表达的技术，是转录后基因沉默的一种研究基因功能的有力工具。

RNAi 的基本原理是利用短双链 RNA 能与细胞内无活性或低活性的酶复合物 Dicer（由核酸内切酶和解旋酶等组成）相结合并使其激活而形成 RNA 诱导的沉默复合物（RISC）的特性，特异性地切割细胞内异常双链 RNA，被切割的双链 RNA 可以产生更多的短双链 RNA，从而形成更多的 RISC，这种放大效应将最终清除所有的靶 RNA（图 19-17）。在这个过程中，短双链 RNA（siRNA）发挥着特异性识别和定位的重要作用。

RNAi 是机体的一种正常防御机制，在正常情况下，细胞内不存在双链 RNA，一旦出现双链 RNA，无论是人工导入的或是病毒 RNA 都会在 RNA 依赖的 RNA 聚合酶（RNA dependent RNA polymerase，RdRp）的催化作用成为双链 RNA，激活 Dicer，从而启动 RNAi，最终清除双链 RNA。

RNAi 技术是研究基因功能的一种公认的有效工具，可以快速评价基因功能和基因无效表型，可以用于筛选药物靶点，并可能产生基于 RNAi 的药物，用于治疗肿瘤、病毒感染等人类疾病。RNAi 技术的基本流程是：①根据目的基因 mRNA 序列确定干涉靶点，一般靶序列应具有 AA（N19）TT or AA（N21）序列特征（其中 N 代表任意核苷酸），同时 GC 比在 50%左右。为了能高效地、特异性地切割目的基因转录产物，靶序列应避开可能的蛋白质结合位点，如 5′非翻译区和 3′非翻译区，以及靠近起始密码的区域，因为这些位点可能是富含调节蛋白结合位点的区域；②设计 siRNA 序列，一般 siRNA 的 3′端应该有 UU 单链（或用 dTdT 单链替代）；③用化学合成法合成 siRNA，或构建特殊的可表达 siRNA 的载体（如用人 U6 启动子启动设计的 siRNA 合成的 pcDNA3-U6）转染细胞，在体内转录合成小发夹 RNA；④将 siRNA 导入细胞中，然后对目标 RNA 含量进行检测，判断 siRNA 的作用。

利用 RNAi 技术寻找药物的作用靶点可以解决以往任何一种药物筛选方法都无法解决的难题，即特异性问题和候选序列问题，而且利用 RNAi 技术可以在疾病复杂的多细胞信

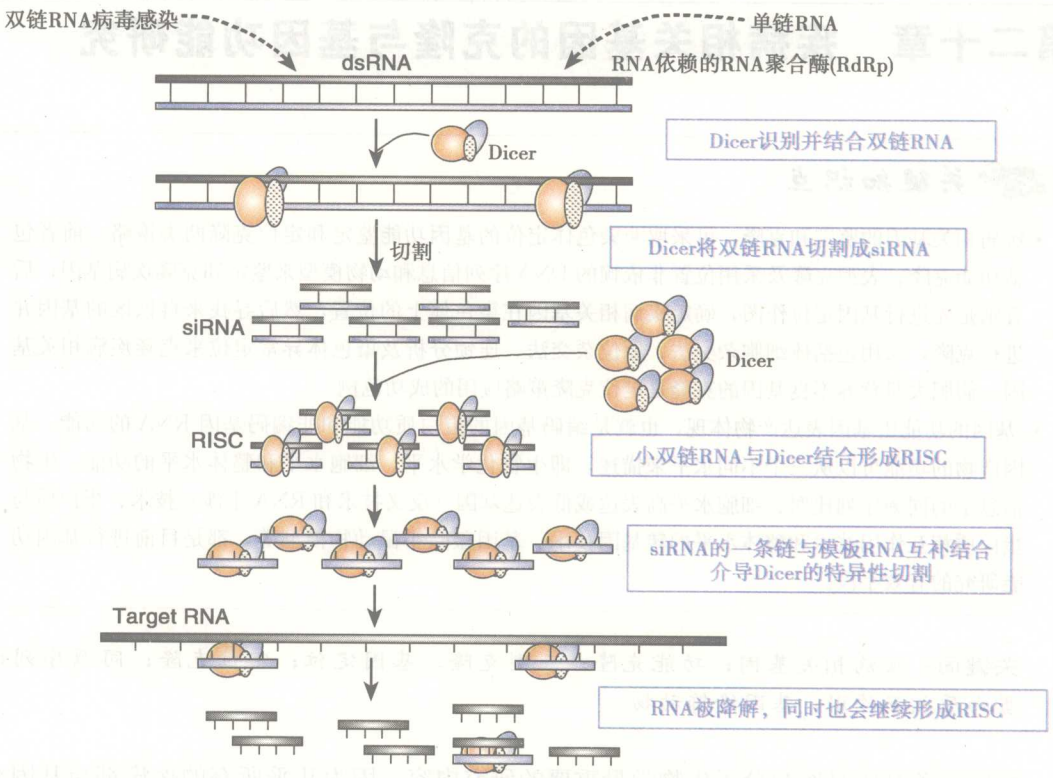


图 19-17　RNA 干涉原理示意图

号途径中发现药物靶点，直接产生核酸和蛋白质候选药物。基于 RNAi 技术的药物产业正在兴起，主要包括抑制基因功能的核酸和蛋白质类药物、多肽、以及 RNAi 本身。而且，目前多家公司正在探索 siRNA 的体内给药途径和方式。

（王丽颖）

第二十章　疾病相关基因的克隆与基因功能研究

关键知识点

- 疾病相关基因的鉴定和克隆，可采取非染色体定位的基因功能鉴定和定位克隆两类策略。前者包括功能克隆、表型克隆及采用位置非依赖的DNA序列信息和动物模型来鉴定和克隆疾病基因；后者则是先进行基因定位作图，确定疾病相关基因在染色体上的位置，然后寻找来自该区的基因并进行克隆，采用包括体细胞杂交法、原位杂交法、连锁分析及染色体异常定位来克隆疾病相关基因。假肥大肌营养不良基因的克隆是定位克隆策略应用的成功范例。
- 基因的功能由基因表达产物体现，也就是编码基因的蛋白质功能和非编码基因RNA的功能。基因产物的功能可以从三个不同水平来描述，即生物化学水平、细胞水平和整体水平的功能。生物信息学的同源序列比对、细胞水平高表达或低表达基因（反义技术和RNA干涉）技术、蛋白质与蛋白质相互作用技术和整体水平的转基因技术、基因敲除小鼠动物模型等，都是目前进行基因功能研究的有效手段。

关键词：疾病相关基因；功能克隆；表型克隆；基因定位；定位克隆；同源序列比对；蛋白质相互作用；基因修饰动物

疾病相关基因是医学分子生物学最重要的研究内容，因为几乎所有的疾病都与基因有关，都是遗传因素和环境因素的相互作用的结果。依据基因变异在疾病发生、发展中的作用，这些基因可分为：①致病基因　如果一种疾病的表型和一个基因型呈现直接对应的因果关系，即该基因结构或表达的异常是导致该病的发生的直接原因，那么该基因就属于致病基因。这类疾病主要是单基因病，即传统的遗传性疾病，环境因素的影响较小。疾病在群体中的发生遵从孟德尔遗传定律，故亦称之为孟德尔遗传病。②疾病易感基因　复杂性疾病，诸如肿瘤、心血管疾病、代谢性疾病、自身免疫性疾病等，环境因素和遗传因素均起着一定作用，表现为2个以上基因的“微效作用”的相加，或者多基因间的相互作用。单一基因的变异仅增加对疾病的易感性，亦称为多基因病。为讨论方便，可笼统地将影响疾病发生发展的基因，均称为疾病相关基因（disease related genes）。

鉴定疾病相关基因，不但可以详尽地了解疾病的病因和发病机制，开发新的诊断和干预技术，而且有利于了解基因的功能，因而一直是生物医学工作者研究的重点。鉴定新基因也具有重大的商业价值，各大生物药物公司纷纷参入这一基因发现的大战中。人类基因组计划的完成为疾病相关基因的发现提供了有利的契机。后基因组计划的主要任务之一就是诠释基因的功能，而克隆疾病相关基因与确定基因功能二者有着密不可分的联系，并可相互促进。

第一节　鉴定克隆疾病相关基因的原则

一、鉴定克隆疾病相关基因的关键是确定疾病表型和基因间的实质联系

疾病表现为症状、体征及其生物化学的异常，可以借助临床查体、实验室检查及其他辅助检查，结合病史而确诊。正确的诊断是鉴定疾病基因的先决条件，在一些复杂性疾病中，有必要进行进一步的分类，确保疾病的同质性，以减少临床的异质性。其次，确定疾

病的遗传因素，即通过孪生子分析，领养分析和同胞罹患率分析等确定遗传因素是否在疾病发病中具有作用及其作用程度。在遗传因素作用较小的疾病中，鉴定并最终克隆疾病相关基因成功的可能性很小。一旦确定了遗传因素在疾病的重要作用，就可进而确定存在于人类基因组中决定疾病表型的基因，定位该基因在基因组中的位置（位点）以及该位点和基因组其他位点的联系。

二、鉴定克隆疾病相关基因需要多学科多途径的综合策略

鉴定克隆疾病相关基因是一项艰巨的系统工程，需要多学科的紧密配合，针对不同疾病采用不同的策略。如图 20-1 所示，这些不同策略和方法结合，互为补充，方可达到最终克隆疾病基因的目的。首先通过对不同疾病家系的连锁分析，可粗略将疾病的基因定位于某个染色体上。这些位点被称之为疾病的位点（locus），此时确切的基因尚不明了。随着对某些疾病发病机制的了解，可以进一步阐明疾病相关的蛋白质的生物化学和细胞生物学特征，以此为切入点，寻找基因结构的异常，确定基因 DNA 序列碱基突变及其导致蛋白质结构或表达异常的分子机制，最终克隆疾病的基因。疾病动物模型对于疾病相关基因的克隆具有重要帮助。对不同的疾病动物模型的研究，可以确定导致实验动物异常表型的基因，进而鉴定人类的同源基因在疾病中的作用。借助生物信息数据库的有关待定基因的信息，亦可极大促进了疾病相关基因克隆的效率。

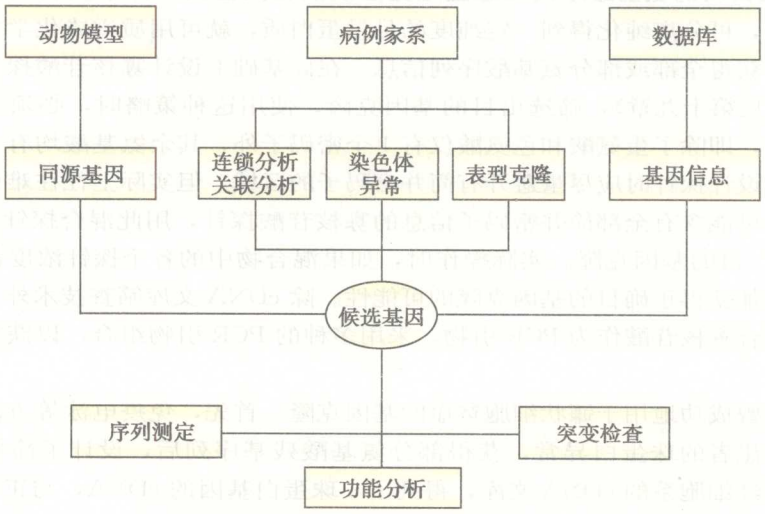


图 20-1　疾病相关基因鉴定克隆策略示意图

三、确定候选基因是多种克隆疾病相关基因方法的交汇

有许多途径能够达到最终鉴定疾病相关基因的目的，这些方法最终将交汇在候选基因上。一旦候选基因被鉴定，即可筛检患者中该基因的突变。候选基因可不依赖其在染色体位置而鉴定，但常用的策略仍然是首先找出候选染色体区域，然后在此区域内鉴定候选基因。人类基因组计划的完成提供了人类所有基因的信息。尽管从众多的候选基因中寻找基因变突仍然是一项耗时的工作，但鉴定候选基因较以往还是相对容易。这是由于位置信息从 2～3 万个的基因降低到候选区域的 10～30 个基因。在候选区域内预测可能的候选基因

并非易事，目前仍然缺乏这样的能力。需要的是大量的重复、逐个排除，最终确定候选基因的突变以及这种突变和疾病的联系。

需要注意的是不要被囊性纤维化基因、糖尿病基因等基因名称所误导。许多人类基因首先是通过研究该基因突变导致的疾病时发现的，故以疾病来命名。但是实际上这些基因均具有正常的生物学功能。

第二节　基因相关疾病克隆的策略和方法

一、不依赖染色体定位的疾病相关基因克隆策略

不依赖染色体定位的疾病相关基因克隆策略包括功能克隆、表型克隆及采用位置非依赖的 DNA 序列信息和动物模型来鉴定和克隆疾病基因。

（一）从已知蛋白质的功能和结构出发克隆疾病基因

在掌握或部分了解基因功能产物蛋白质的基础上，鉴定蛋白质相关基因，进而克隆该基因的方法，称之为功能克隆（functional cloning），这是相对于利用基因位置克隆基因的定位克隆而言的。该方法采用的是从蛋白质到 DNA 的研究路线，针对的是一些对影响疾病的功能蛋白具有一定了解的疾病，如血红蛋白病、苯酮酸尿症等出生缺陷引起的分子病可以采用这个方法定位和克隆疾病基因。

1. 依据蛋白质的氨基酸序列信息鉴定克隆疾病相关基因　如果疾病相关的蛋白质在体内表达丰富，可分离纯化得到一定纯度的足量蛋白质，就可用质谱或化学方法进行氨基酸序列分析，获得全部或部分氨基酸序列信息。在此基础上设计寡核苷酸探针，用于筛查 cDNA 文库（见第十九章），筛选出目的基因克隆。使用这种策略时，必须考虑到密码子的简并性特点，即除了蛋氨酸和色氨酸仅有 1 个密码子外，其余氨基酸均有 2 个或 2 个以上的密码子。设计探针时应尽量避开有简并密码子的区域，但实际上往往难以做到。为此可以设计 1 套可能含有全部简并密码子信息的寡核苷酸探针，用此混合探针去筛查 cDNA 文库，“钓出”目的基因克隆。实际操作时，如果混合物中的各个探针浓度都保持在低水平，可明显增加获得正确目的基因克隆的可能性。除 cDNA 文库筛查技术外，目前还可采用部分简并混合寡核苷酸作为 PCR 引物，采用多种的 PCR 引物组合，以获得候选基因的 PCR 产物。

上述方法曾成功地用于镰状细胞贫血的基因克隆。首先，免疫电泳等方法已经显示出镰状细胞贫血患者的珠蛋白异常，获得部分氨基酸残基序列后，设计了简并寡核苷酸探针，筛选有核红细胞系的 cDNA 文库，得到了 α 珠蛋白基因的 cDNA，与正常人的 cDNA 比较，发现了 α 珠蛋白基因变异。进而找出 cDNA 探针与染色体 DNA 序列间的同源互补关系，将人的 α 珠蛋白基因定位于第 16 号染色体上。

2. 用蛋白质的特异性抗体鉴定疾病基因　有些疾病相关的蛋白质在体内含量很低，难以纯化得到足够纯度的蛋白质用于氨基酸序列测定。但是少量低纯度的蛋白质仍可用于免疫动物获得特异性的抗体，用以鉴定基因。获得的抗体一方面可用于直接结合正在翻译过程中的新生肽链，此时会获得同时结合在核糖体上的 mRNA 分子，最终克隆未知基因；另外，特异性抗体也可用来筛查可表达的 cDNA 文库，筛选出可与该抗体反应的表达蛋白质的阳性克隆，亦可获得候选基因。

功能克隆仍然是单基因疾病基因克隆的常用策略。其缺点是特异功能蛋白质的确认、鉴定及其纯化都相当困难，微量表达的基因产物在研究中难以获得，因而几乎不能用于多基因疾病的基因分离。

（二）从疾病的表型差异出发克隆疾病相关基因

表型克隆（phenotype cloning）是疾病相关基因克隆领域中一个新的策略。该策略的原理是基于对疾病表型和基因结构或基因表达的特征联系已经有所认识的基础上来分离鉴定疾病相关基因。依据 DNA 或 mRNA 的改变与疾病表型的关系，可有几种策略：①从疾病的表型出发，比较患者基因组 DNA 与正常人基因组 DNA 的不同，直接对产生变异的 DNA 片段进行克隆，而不需要基因的染色体位置或基因产物的其他信息。例如，在一些遗传性神经系统疾病中，患者基因组中含有的三联重复序列的拷贝数可发生改变，并随世代的传递而扩大，称为基因的动态突变（见第十章）。此时，采用基因组错配筛选（genome mismatch scanning）、代表性差异分析（representative difference analysis，RDA）等技术即可检测患者的 DNA 是否有三联重复序列的拷贝数增加，从而确定患病原因；②如果高度怀疑某种疾病是由于某个特殊的已知基因所致，可通过比较患者和正常对照间该基因表达的差异来确定该基因是否为该疾病相关基因。常用分析方法有 Northern 印迹法、RNA 酶保护试验、RT-PCR 及实时定量 RT-PCR 等（见第十九章）；③如果基因是完全未知的，也可通过比较疾病和正常组织中的所有 mRNA 的表达种类和含量间的差异，从而克隆疾病相关基因。这种差异可能源于基因结构改变，也可能源于表达调控机制的改变。常用的技术有 mRNA 差异显示（mRNA differential display，mRNA-DD）、抑制消减杂交（suppressive substractive hybridization，SSH）、基因表达的系列分析（serial analysis of gene expression，SAGE）、cDNA 微阵列（cDNA microarray）和基因鉴定集成法（integrated procedure for gene identification）等。这里分别介绍 RDA 和 mRNA-DD 技术。

1. RDA 技术是建立在核酸差异杂交技术基础上的 PCR 技术　RDA 是通过对正常和疾病组织的 cDNA 差异片段（即代表性片段）的扩增，从而使其被检测和捕获的技术。基本原理是，首先用 PCR 方法从拟比较的疾病和正常组织获得足够量的 DNA 或 cDNA 片段；然后进行差异杂交，杂交后再用不同引物进行第二次 PCR 反应；在第二次 PCR 反应中，只有两个样品中结构或表达量有差异的 DNA 片段可以得到扩增（图 20-2）。其基本步骤是：①DNA 片段制备：分别提取正常人基因组 DNA（检测 DNA）和患者基因组 DNA（驱动 DNA），用限制性内切酶消化 DNA，获得长度在 150～1 000 bp 之间的片段；②获

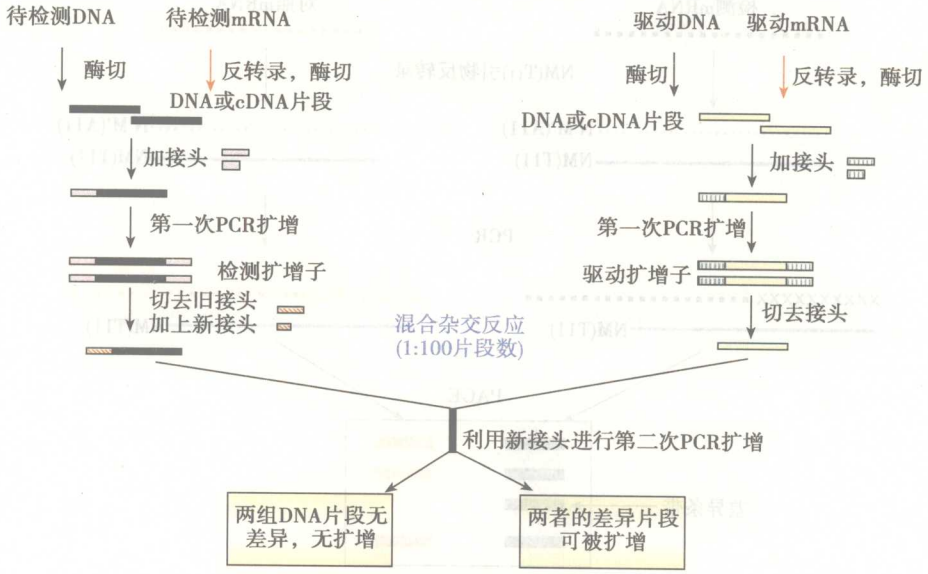


图 20-2　RDA 技术原理和基本过程示意图

得扩增子：在两组的所有 DNA 片段上加上接头，以接头的互补序列为引物，进行第一步 PCR 扩增，所获得扩增产物称扩增子（amplicon）；③更换接头：切去所有扩增子的接头，仅在检测扩增子上加上新的接头；④筛选扩增产物：按 1∶100 的比例混合检测扩增子和驱动扩增子，进行液相杂交。取少量杂交反应物为模板，以新的接头为引物再进行第二次 PCR 扩增，即可筛选出两组 DNA 样品间的差异片段。

检测 DNA 和驱动 DNA 间片段在第二次 PCR 反应中依据两者间是否有差异，主要可以出现两种情况：①两组间相同的 DNA 片段不会得到大量扩增。这是因为在杂交反应中，驱动 DNA 片段的数目远大于检测 DNA，将优先结合检测 DNA，使得检测 DNA 分子间几乎没有机会形成同源复性双链。利用新接头进行的二次 PCR 反应过程将不会有扩增产物。②两者的差异片段可以被扩增。如果检测 DNA 中存在的某一片段在驱动 DNA 中缺失，或由于突变而失去了互补结合能力，在杂交反应中就不存在来自驱动 DNA 中的同源片段的竞争，检测 DNA 自身可以发生复性，且由于复性的双链 DNA 两端都具有新接头，因而可以实现 PCR 的大量扩增。该片段即为候选的疾病相关 DNA 序列。虽然反应中无差异片段还会存在一些机会被扩增，但产物量较小，可以被区分。

RDA 也用于 mRNA 差异表达基因的克隆上，只是需要先将 mRNA 反转录成 cDNA 片段。RDA 技术对正常和异常的 DNA 片段区分能力强、富集效率高、对起始材料要求低，利用 RDA 人们已经发现了多个疾病相关新基因。

2. mRNA-DD 是 RT-PCR 技术和聚丙酰胺凝胶电泳技术的结合　mRNA-DD 又称为差异显示反转录 PCR（mRNA differential display reverse transcription PCR，DDRT-PCR）方法。该法利用可以扩增所有哺乳类生物 mRNA 的几条 5′随机引物和几条 3′锚定引物组合，用 PCR 的方法扩增正常人和患病个体的相应组织的 cDNA。用聚丙酰胺凝胶电泳分离扩增产物，比较两组间产物的差异（图 20-3）。依据理论计算，该方法所设计的组合引物可以与所有 mRNA 的 poly（A）尾匹配，因而对于种类和含量相同的 cDNA 样品，PCR 产物的种类多少和分布应该是完全一样的。如果在正常和患者的 cDNA 标本中扩增出一些不同长度的 cDNA 片段，它们所代表的 cDNA 就有可能与疾病状态相关。这一方法的优点在于所需 mRNA 量少、较快速、可以同时显示多种生物性状的差异及可以同时获

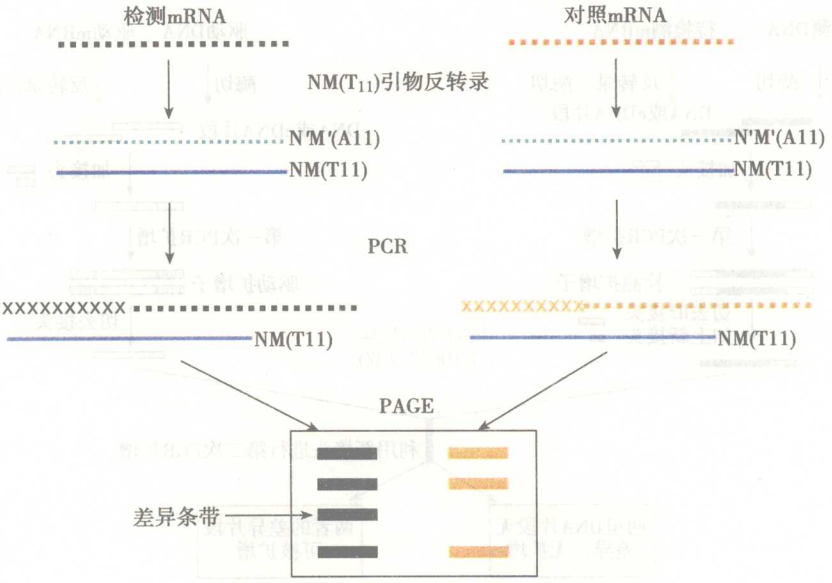


图 20-3　mRNA-DD 技术原理和基本过程示意图

得高表达和低表达的基因等。这种方法同时也存在许多严重的缺陷，如假阳性率高达70%、获得的片段太短等，很难直接判断其功能和意义。尽管有上述缺陷，但因其步骤较简单，可获得较大量信息，在实际工作中该法应用仍较多。

（三）采用动物模型鉴定克隆疾病相关基因

人类的许多种疾病，都有相应的动物模型。如果动物某种表型的突变基因定位于染色体的某一部位，而具有相似人类疾病表型的基因很有可能存在于人染色体的同源部位。另外，当疾病基因在动物模型上已完成鉴定，还可以采用荧光原位杂交来定位分离人的同源基因。肥胖相关的瘦素（leptin）基因的克隆就是一个成功例证，对突变的肥胖近交系小鼠通过定位克隆分离得到了位于小鼠6号染色体的瘦素基因，依据小鼠瘦素基因侧翼标记，将人的瘦素基因定位于人染色体7q31区。小鼠和人的瘦素基因有84%的同源性，编码一167氨基酸的分泌蛋白-瘦素，其主要功能是控制食物的摄入，促进能量的消耗。肥胖小鼠和一些遗传性肥胖症患者均具有该基因的缺损，导致基因功能丧失。

二、定位克隆是后基因组时代的常用策略

仅根据疾病基因在染色体的大体位置，鉴定克隆疾病相关基因，称之为定位克隆（positional cloning）。定位克隆的起点是基因定位，即确定疾病相关基因在染色体上的位置，然后根据这一位置信息，应用DNA标记将经典的遗传学信息转换为遗传标记所代表的特定基因组区域，再以相关基因组区域的相连重叠群（contig）筛选结构基因，最后比较患者和正常人这些基因的差异，确定基因和疾病的关系。人类基因组计划后所进行的定位候选克隆，是将疾病相关位点定位于某一染色体区域后，根据该区域的基因、EST或模式生物对应的同源区的已知基因等有关信息，直接进行基因突变筛查，通过多次重复，最终确定疾病相关基因。

（一）从家系分析出发，在细胞、染色体和分子水平进行基因定位

基因定位（gene location）是基因分离和克隆的基础，目的是确定基因在染色体的位置以及基因在染色体上的线性排列顺序和距离。可从家系分析、细胞、染色体和分子水平等几个层次进行基因定位，由于使用手段的不同可派生出多种方法，不同方法又可联合使用，相互补充。

1. 体细胞杂交法是通过融合细胞的筛查定位基因　体细胞杂交（somatic cell hybridization）又称细胞融合（cell fusion），是将来源不同的两种细胞融合成一个新细胞。大多数体细胞杂交是用人的细胞与小鼠、大鼠或仓鼠的体细胞进行杂交。这种新产生的融合细胞称为杂种细胞（hybrid cell），含有双亲不同的染色体。杂种细胞有一个重要的特点是在其繁殖传代过程中出现保留啮齿类一方染色体而人类染色体逐渐丢失，最后只剩一条或几条，其原因至今不明。Miller等运用体细胞杂交，结合杂种细胞的特征，证明杂种细胞的存活需要胸苷激酶。含有人17号染色体的杂种细胞在特殊的培养基中，都因有TK活性而存活，反之则死亡，从而推断TK基因定位于第17号染色体上。利用这一方法定位了许多人的基因。

2. 染色体原位杂交是在细胞水平定位基因的常用方法　染色体原位杂交（in situ hybridization）是核酸分子杂交技术在基因定位中的应用，也是一种直接进行基因定位的方法。其主要步骤是获得组织培养的分裂中期细胞，将染色体DNA变性，与带有标记的互补DNA探针杂交，显影后可将基因定位于某染色体及染色体的某一区段。如果用荧光染料标记探针，即为荧光原位杂交（fluorescence in Situ hybridization，FISH）。1978年首次用α及β珠蛋白基因的cDNA为探针，与各种不同的人/鼠杂种细胞进行杂交，从而将

人α及β珠蛋白基因分别定位于第16号和第11号染色体上。这种染色体原位杂交技术特别适用于那些不转录的重复序列，这些重复序列很难用其他方法进行基因定位。如利用原位杂交技术将卫星DNA定位于染色体的着丝粒和端粒附近。

3. 染色体异常有时可提供疾病基因定位的替代方法　从基因定位克隆的角度来看，对于任何已知与染色体异常（chromosome abnormalities）直接相关的疾病来说，染色体的异常本身就成为疾病定位克隆的一个绝好的位置信息。染色体的异常有时可替代连锁分析，用于定位疾病基因。在一些散发性、严重的显性遗传病，染色体变异分析是获得候选基因的唯一方法。有时可直接获得基因的正确位置，而无需进行连锁分析，例如染色体的平衡易位和倒位等。诸如多囊肾、巨肠症、假肥大型肌营养不良基因的定位在很大程度上借助于染色体的异常核型表现。

如果细胞学观察的染色体异常与某一基因所表达的异常同时出现，即可将该基因定位于这一染色体的异常区域内。例如对一具有6号染色体臂间倒位的家系分析表现，凡是有此倒位者，同时也都有某一*HLA*等位基因的表达，而家族中无此倒位者，也无该等位基因的表达，因此将人*HLA*基因定于6号染色体短臂的远侧区。

染色体非整倍体分析中，可通过基因剂量法进行基因定位。在先天愚型（核型47，+21）的病人中过氧化物歧化酶-1的活性比正常人高1.5倍，因此将该酶基因定位于21号染色体上。但是并非所有基因的拷贝数都有明显的剂量效应作用。

4. 连锁分析是定位疾病未知基因的常用方法　基因定位的连锁分析（linkage analysis）是根据基因在染色体上呈直线排列，不同基因相互连锁成连锁群的原理，即应用被定位的基因与同一染色体上另一基因或遗传标记相连锁的特点进行定位。如果待定基因与标记基因呈连锁遗传，即可推断待定基因与标记基因处于同一染色体上，并且依据和多个标记基因连锁的程度（用两者间的重组率度量），确定待定基因在染色体的排列顺序以及和标记基因间的遗传距离。例如已知血型基因Xs定位于X染色体上，普通鱼鳞病和眼白化病基因与其连锁，因此判定这两个基因也在X染色体上，计算患者子代的重组率，即可确定这些基因间的相对距离。

在基因定位克隆中，常用一些分子遗传标记进行全基因组的连锁分析和关联研究（见第二十一章），极大地提高了基因克隆的效率。

（二）定位克隆的基本程序包括三大步骤

1. 尽可能精确染色体上的候选区域　定位克隆困难与否取决于染色体候选区域的大小。为此要尽可能的缩窄疾病基因在染色体的候选区域。在单基因疾病基因的遗传制图时，需要选择更多的遗传标记，寻出遗传距离最近的标记，增加更多的家系、建立所有个体的单倍体型等，以增加发现重组机会，结合寻找更多连锁不平衡，精确疾病相关基因的候选区域。

2. 构建目的区域的物理图谱　由于人类基因组研究的发展，各种DNA分子水平上物理手段的建立，已经使得疾病相关基因的克隆变得较为容易，现在已无需建立DNA重叠群，仅需从人类基因组序列数据库下载即可，但使用前要仔细检查重叠的拼装是否正确。

3. 疾病相关基因的确定　在得到区域性很窄的DNA重叠群克隆后，可以使用多种方法对变异位点进行确定。常用的方法主要有：①cDNA直接筛选法：如果知道该种疾病发生的特异组织，还可以将该组织中的cDNA直接与得到的克隆杂交，筛选出此区域内的特异表达基因，再对这些基因作进一步分析；②候选基因克隆法：对该区域中的已知基因位点进行测序，比较变异情况，确定变异位点；③对克隆直接进行序列分析。

（三）假肥大型肌营养不良基因的克隆是定位克隆的成功例证

采用定位克隆策略鉴定的第一个疾病相关基因是X连锁慢性肉芽肿病基因。而假肥大

型肌营养不良（Duchenne muscular dystrophy，DMD）基因的成功克隆，更彰显了基因定位克隆的优势。这项工作主要分两个阶段。首先，根据患病女性X染色体与第21号常染色体的易位，以及男患儿发生小的Xp21.2缺失并伴发三种其他X连锁隐性遗传病，再运用RFLP连锁分析将DMD基因定位于Xp21。然后，分别克隆得到了基因的2个不同的片段，分别命名为XJ系列探针和pERT87系列探针，根据两片段的比较，证明DMD基因约为2 300 kbp，占X染色体的1%以上，该基因编码肌营养不良蛋白（dystrophin），影响横纹肌和心肌的结构和收缩功能。

三、利用生物信息数据库有助于疾病相关基因克隆的研究

人类基因组计划和多种模式生物基因组测序的完成，生物信息学的发展，计算机软件的开发应用和互联网的普及，人们通过已获得的序列与数据库中核酸序列及蛋白质序列进行同源性比较，或对数据库中不同物种间的序列比较分析、拼接，预测新的全长基因等，进而通过实验证实，从组织细胞中，克隆该基因，这就是所谓的电子克隆（in silico cloning）。

人类新基因克隆大都是从同源表达序列标签（EST）分析开始的。应用同源比较，在人类EST数据库中，识别和拼接与已知基因高度同源的人类新基因的方法包括：①以已知基因cDNA序列对EST数据库进行搜索分析（Basic Local Alignment Search Tool，BLAST），找出与已知基因cDNA序列高度同源的EST；②用Seqlab的Fragment Assembly软件构建重叠群，并找出重叠的一致序列；③比较各重叠群的一致序列与已知基因关系；④对编码区蛋白质序列进行比较，并与已知基因的蛋白质的功能域进行比较分析，推测新基因的功能；⑤用新基因序列或EST序列对序列标签位点（sequence-tagged site，STS）数据库进行BLAST分析，如果某一EST（非重复序列）与某一种STS有重叠，那么，STS的定位即确定了新基因的定位。

电子克隆充分利用网络资源，可大大提高克隆新基因的速度和效率。由于数据库的不完善、错误信息的存在及分析软件的缺陷，电子克隆往往难以真正地克隆基因，而是一种电子辅助克隆。

第三节　疾病相关基因的功能研究

基因的功能实际是基因产物的功能，也就是编码基因的蛋白质功能和非编码基因RNA的功能，前者是本节讨论的重点。非编码RNA，如MicroRNA的功能，最近亦引起广泛的关注（见第六章）。基因产物的功能可以从三个不同水平来描述，即生物化学水平，主要描述基因产物参与了何种生化过程，如属于激酶、转录因子等；细胞水平，主要论述基因产物在细胞内的定位和参与的生物学途径，例如，某蛋白质定位于核内，参与DNA的修复过程，有可能并不了解确切的生物化学功能；整体水平的功能，主要包括基团表达的时空性及基因在疾病中的作用。

为研究基因的功能，有必要获得尽可能多的有关该基因的信息。目前生物信息学方法可提供很大的帮助。对基因产物不同水平的研究，需要采用不同的研究手段。例如采用生物信息学序列比对，预测基因功能；通过细胞水平上的基因高表达或基因沉默，观察细胞的功能改变；研究蛋白质与蛋白质或DNA的相互作用，了解基因产物的生物学途径；构建转基因或基因敲除动物，在整体水平研究基因功能。

克隆了疾病相关基因，往往并不了解基因的正常功能。对人类基因功能的研究是后基

因组时代生命科学领域中的重大课题。通过基因功能的研究可以阐明人体细胞的增殖、分化、通讯、衰老和死亡的分子机制，确定人类疾病发生发展及转归的机制，进而研发新的诊断技术及治疗干预措施，提供药物开发的靶标分子。

一、基因比对及功能诠释

在以往的研究中，已经对大量的基因功能产物的功能有详尽的了解，获得的足够多的信息，建立了共享资源数据库，其中最为著名就是美国的 GenBank。这些数据库是进行基因序列比对，诠释基因功能的基础。依据分子进化的理论，核酸或氨基酸序列相似的基因，应表现出类似的功能。这些序列相似的基因称之为同源基因。基本的方法是基因序列比对分析。

两个或多个符号序列按字母比较、尽可能确切地反映它们之间的相似和差异，称为序列比对，又称序列联配。通常所讲的序列比对包括核酸序列比对和蛋白序列比对。根据每次参与比对的序列数又可分为双序列比对和多序列比对。

常用的两大双序列比对工具是 BLAST 和 FASTA，分别由美国国立生物技术信息中心（NCBI）和欧洲生物信息学研究所（EBI）开发和维护。NCBI 的主页是 http：//www. ncbi. nlm. nih. gov/；EBI 的主页是 http：//www. ebi. ac. uk/。两个数据库均有详细的序列比对数据，且有方便的对话框式操作和使用说明。任何一个新基因的功能研究的第一步就是在这些数据库中进行序列比对。

二、利用工程细胞研究基因功能

在细胞水平研究基因功能，除了观察细胞在实验条件下该基因表达的改变，更重要的是人为地导入外源基因或干预正常基因的表达，以观察细胞生物学的改变。

（一）采用基因重组技术建立基因高表达工程细胞系

采用基因重组技术，将外源基因导入宿主细胞，使其表达目的基因，进而观察细胞的生物学特征改变。根据外源基因在宿主细胞表达的持续时间，可区分为瞬时转染和稳定转染细胞。稳定转染细胞是一种最常用的细胞水平的转基因模型，外源基因通过转基因过程插入到细胞的染色体中，或以游离基因（episome）的形式存在于细胞中稳定表达。稳定表达外源蛋白的转染细胞可以用作基因功能研究的细胞模型，也可以利用这种方法获得外源基因产物，目前被广泛应用于细胞模型的建立。

（二）基因沉默技术抑制特异基因的表达

在细胞水平沉默待研究基因的表达，进而观察基因沉默后细胞的生物学行为改变，是基因功能研究的有力工具。可以利用 RNA 干扰技术、反义技术等来实现对特定基因表达的抑制。

三、研究蛋白质分子间的相互作用，通过了解生物学途径确定基因功能

在一些情况下，即使对基因的序列、结构和表达模式有清楚的认识，但仍然难以阐明基因表达蛋白质的功能。此时通过研究该蛋白质与已知功能蛋白质的相互作用，无疑将非常有助于对该蛋白质功能的了解。例如，一未知功能的新蛋白质与一参与 RNA 剪接的蛋白质存在相互作用，极有可能该蛋白质也参加了 RNA 的剪接过程。研究蛋白质相互作用的方法包括遗传学、生物化学和物理方法。遗传学方法仅用于如果蝇、酵母等模式生物，

而生化和物理的方法可直接用于人类细胞。常用的生化方法是亲和层析和免疫共沉淀。常用的高通量筛查蛋白质间相互作用的方法是酵母双杂交技术和噬菌体展示技术。

酵母双杂交系统（yeast two-hybrid-system）是基于真核细胞转录激活因子的结构和活性特点，在酵母中研究蛋白质相互作用的高通量检测系统。具体原理是利用了酵母细胞的一个转录激活因子的DNA结合结构域（binding domain，BD）和转录激活结构域（activation domain，AD）。这两个结构域对维持转录激活因子的功能缺一不可。酵母蛋白GAL4是一个典型的转录激活因子，可活化酵母半乳糖苷酶。将待进行相互作用分析的编码蛋白X的基因序列与GAL4的BD部分融合，形成诱饵基因；编码相互作用候选分子蛋白Y的基因序列与GAL4的AD基因融合，形成猎物基因。在GAL4调控的下游融合有报告基因，用于筛查。如果编码蛋白X和Y相互作用，就可使BD和AD靠近形成一有效的转录因子，激活报告基因的表达；反之，如无相互作用，就不能激活报告基因的表达。如果以某一组织的cDNA文库做为猎物基因，即可进行高通量的蛋白质相互作用筛查，由于存在一定的假阳性，对筛选出的候选克隆还需要用其他方法确定这种相互作用。

噬菌体展示（phage display）是将外源性DNA插入到噬菌体的衣壳蛋白基因中的一种表达克隆技术。重组基因以融合蛋白的形式，参入病毒颗粒中，展示在噬菌体的表面。融合噬菌体可与展示在噬菌体表面的外源蛋白相互作用的蛋白结合，以筛查蛋白质的相互作用。噬菌体展示的缺点是本身是一种体外分析系统，另外，只有短肽可以展示在噬菌体的表面。噬菌体展示也可用于抗体工程和普通的蛋白质工程。

四、基因修饰动物是整体研究基因功能的重要方法

尽管在体外对基因功能的研究可提供大量信息，但是只有在整体内的研究才能真实反映该基因在体内的真正作用。在实验动物，特别是小鼠体内进行相关的基因操作，通过受精卵或胚胎干细胞中表达外源性基因、敲入或敲除已有基因，获得基因修饰动物品系，已成为在体研究基因功能的重要手段，常用的有转基因动物和基因敲除动物，其具体原理和用途见第十九章。

（吕社民）

第二十一章　基因诊断

关键知识点

- 基因诊断的基本概念：基因诊断主要是针对 DNA 分子的遗传分析技术，包含定性和定量两类分析方法。基因诊断的基本操作流程包括核酸抽提、目的序列扩增、分子杂交和信号检测，其中 PCR 扩增和分子杂交是现代基因诊断技术的基本方法。基因诊断已成为临床实验医学的一个重要组成部分。
- 基因诊断的技术原理：遗传病基因诊断的基本方法学包括连锁分析和直接诊断。用于连锁分析的遗传标志物主要有 RFLP、STR 和 SNP。用于遗传分析的代表性直接诊断技术有：用于 DNA 分析的 Southern 印迹杂交法、用于检测点突变的 ASO 分子杂交法以及用于识别 STR 序列的 PCR 直接扩增法等。
- 遗传检测技术的医学应用：目前在临床医学诊断、遗传病预防，以及法医学个体认定等领域有了广泛的应用，在临床应用上，人类遗传病，尤其是单基因病的确诊及其症状前诊断是目前遗传诊断技术应用领域的成功典范。未来在应用于多基因常见病的发病风险预测，个体化用药指导和疗效评价等方面将显示出巨大的应用潜力。在遗传病预防方面，产前诊断和植入前遗传学诊断技术可作为干预出生缺陷发生的重要工具。

关键词： 基因诊断；分子杂交；连锁分析；直接诊断

基因诊断的概念源于对人类遗传病的认识，其最早的医学实践始于 1978 年美籍华裔科学家简悦威（YW Kan）博士的研究。他利用与致病基因相关联的 DNA 多态性分析技术，成功地进行了镰状细胞贫血症这一遗传性血红蛋白病的特异性产前基因诊断，从而开创了基因诊断技术在临床应用的新时代。传统的疾病诊断都是以疾病的表型改变为依据，随着分子医学知识的不断积累，人类对疾病的认识更注重寻找和发现导致表型发生的原因，即希望依据人体遗传物质的改变或基因变异来诊断疾病，特别是人类单基因遗传病的诊断，基因诊断不仅是临床确诊的必要步骤，也是其根本依据。可以说人类绝大多数疾病的发生和发展，或多或少与自身的基因相关。相信随着像高血压这样的多基因相关常见病分子病因的阐明，基因诊断将得以更广泛的应用，并推动个体化医学的发展。

近 30 年来，克隆定位人类单基因遗传病的基因一直是分子遗传学研究的主题，人类基因组计划全序列图的完成大大推动了这项研究的进程，遗传病的基因诊断是基础研究在医学上应用的最重要的成果之一。有关人类疾病的分子医学知识的迅速积累和现代分子生物学技术的发展，使基因诊断已成为临床遗传学这一新兴学科发展的重要支撑点，此外，基因诊断技术还在肿瘤学、法医学、组织病理学、细菌学和病毒学等领域有着广泛的应用。就基因诊断的应用对象而言，疾病的基因诊断是一个有广泛含义的名词，如传染病病原体的诊断也属基因诊断。本章将主要围绕人类遗传病的基因诊断来展开阐述，着重介绍遗传缺陷诊断的技术原理及其主要应用，这些知识也是理解基因诊断在其他方面应用的基础。

第一节 基因诊断的概念

一、DNA是基因诊断的主要对象

从广义上说，凡是用分子生物学技术对生物体的DNA序列及其产物（如mRNA和蛋白质）进行的定性、定量分析，都称为分子诊断（molecular diagnosis）。通常又将针对DNA的分子诊断称为DNA诊断或基因诊断（gene diagnosis）。从技术角度讲，目前的分子诊断方法主要是针对DNA分子的，涉及到功能分析时，需定量检测RNA（主要是mRNA）靶分子。由于分析技术相对复杂，目前很少将蛋白质分子作为常规分析对象。DNA和mRNA分子都是遗传信息的载体，人类遗传病的基因诊断就是分析这些遗传信息分子的序列，从而在分子水平上确定疾病的病因所在。由于基因变异是导致基因功能改变并产生疾病的根本原因，同时，变异的化学基础又几乎是独有的，故基因诊断是在源头上识别基因正常与否，具有高度的特异性。在人类遗传病的临床诊断中，还包括以染色体分析为重点的细胞遗传学诊断和以蛋白质为主要对象的生化遗传学诊断，本章不予讨论。

二、基因诊断包括定性和定量分析

遗传病基因诊断的前提是疾病表型与基因型关系已被阐明。开展遗传病的基因诊断必须了解相关基因的染色体定位、基因克隆和功能分析等一些基础知识。人类遗传病的表型是由个体的基因型决定的，故对遗传病的诊断也可理解为进行个体的基因分型。就基因诊断的一般内容而言，包括检测个体的基因序列的特征、基因突变分析、测定基因的拷贝数、分析基因表达产物以及检测是否存在外源基因等。在进行不同内容的基因诊断时，需选择适宜的技术，基因诊断技术大致可分为定性和定量分析两类技术。在遗传病诊断时，基因分型和检测基因突变一般采用定性分析，测定基因（染色体）拷贝数及基因表达产物则属于定量分析。在检测外源感染性病原体（基因）时，定性分析可诊断其在人体存在与否，而定量分析则可确定其拷贝数（含量）。

三、基因诊断操作的四个基本步骤

基因诊断的基本技术流程包括样品的核酸抽提、目的序列的扩增、分子杂交和信号检测。样品的核酸抽提是基因诊断操作的第一步。临床上常用的样品主要包括血液、组织块、羊水和绒毛等，其他样品还有精液、毛发、唾液和尿液等。在选择被测样品时，可根据材料来源和分析目的提取其基因组DNA或RNA，后者可经反转录形成cDNA。RNA提取须采用新鲜样品才能获取。在开展胎儿DNA诊断时，除传统的羊水、绒毛和脐带血样品外，从母亲外周血中提取胎儿细胞或胎儿DNA的先进技术已经初步应用于临床实践，该技术代表了未来产前诊断的发展方向。

待测目标序列的扩增是基因诊断操作的第二步。无论是何种样品来源的核酸，一般均须经过PCR特异性扩增制备足够拷贝数的待测靶DNA片段，供进一步进行特异性分析时使用。有关PCR技术的原理见本书第十九章。在某些情况下，PCR方法本身既是获取样品的手段，也是分析某一特定序列是否存在的检测方法。

分子杂交是基因诊断操作的第三步。确认DNA中是否存在某一特定序列的最基本的

方法，就是核酸分子杂交，又称为核酸分子探针技术。该技术是利用碱基互补配对的原理，以已知的核酸片段作为探针，检测样本中是否存在及有多少与其互补的核酸序列。目前基因诊断常用的核酸分子杂交技术（见第十九章）主要有：Southern 印迹杂交、Northern 印迹杂交、点杂交、等位基因特异性寡核苷酸分子杂交、原位杂交和 DNA 芯片技术等。就人类遗传病的诊断而言，其主要目标是确定变异序列的存在，也可称为突变分析（mutation analysis）。这一步骤还可采用直接 DNA 序列分析和其他 DNA 结构分析方法。DNA 突变分析有许多不同原理的技术方法，不管是否进行核酸分子杂交操作，这一分析步骤可理解为基于特异性 DNA 序列的检测。

信号检测是基因诊断操作的最后一步，即对上述步骤得到的反应结果（DNA 分子是否存在或其拷贝数多少）加以检测。通常有三种分析方法，一是直接检测标记在核酸分子上的示踪物（如放射性核素^{32}P、绿色荧光素 SYBR Green 等）；二是采用酶-底物体系介导的颜色反应检测标记在核酸分子上的示踪物（如生物素和地高辛等）；三是利用核酸分子在 280 nm 波长有强吸收的特性测定其紫外吸光度值。借助相应的检测仪器或通过肉眼观测颜色变化或荧光信号可实现核酸分子的定性或定量检测。

第二节　基因诊断技术的方法学

在基因水平诊断疾病，尤其是人类遗传病的基因诊断，需根据致病基因的结构、突变谱特征或基因的连锁关系特性等，设计建立针对目标人群靶基因的分子诊断方法。本节以遗传病的基因诊断为例，重点介绍连锁分析和直接诊断这两种应用于遗传病突变分析的基本技术的原理和方法。

一、利用遗传标志连锁分析间接检测遗传病

（一）用于连锁分析的遗传标志

连锁是指同一条染色体上位置相邻的基因，常被一起遗传而没有发生重组。连锁分析（linkage analysis）就是利用与致病基因相连的某些基因座位作为遗传标志，通过鉴定遗传标志的存在而判断个体是否带有致病基因。连锁分析的优点是无需更多地了解致病基因结构及其分子机制，适用于大多数由未知的基因缺陷引起的遗传性疾病的基因诊断。由于连锁分析没有直接检测致病的基因突变，又称为间接诊断（indirect diagnosis），它是在不了解致病基因的条件下，进行遗传病基因诊断的一种有效手段。自 20 世纪 80 年代初期以来，基于限制性片段长度多态性（restriction fragment length polymorphism，RFLP）的连锁分析技术，曾成功地用于包括镰状细胞贫血症、β-地中海贫血、亨廷顿病和囊性纤维化等多种人类单基因遗传病在内的基因诊断和遗传分析。

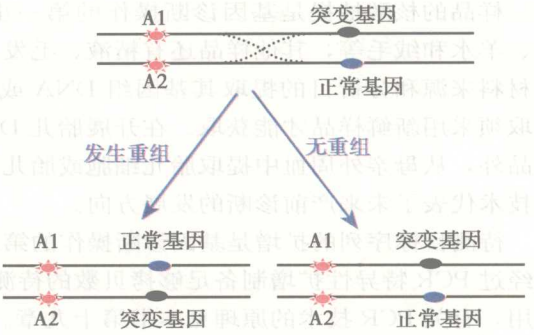


图 21-1　标记基因 A1 和 A2 与致病基因之间的重组改变连锁关系示意图

用于进行连锁分析的遗传标志，需要具备三个特点：①不同个体之间存在高度的多态性。多态性是指在一个特定的基因座位上，存在两个或两个以上的等位基因，而且其中的任何一个在人群中的频率大于 1%，大于 10%的人群常见多态位点

的应用价值更大；②需要获得足够数量的家系成员样本，以区分和确定标志基因与致病基因之间的连锁关系；③标志基因与致病基因相连锁，而且两者之间的距离越小越好，以尽量排除减数分裂中染色体在标志基因与致病基因之间出现重组或交换带来的错误诊断（图 21-1）。为满足在尽可能多的染色体位点进行遗传病的连锁分析，就要求标志基因的数量要足够多，最好能够高密度地遍布整个基因组。根据上述条件，目前用于诊断的遗传标志主要包括两种形式的 DNA 变异：单核苷酸变异和短串联重复序列变异（见第十章）。

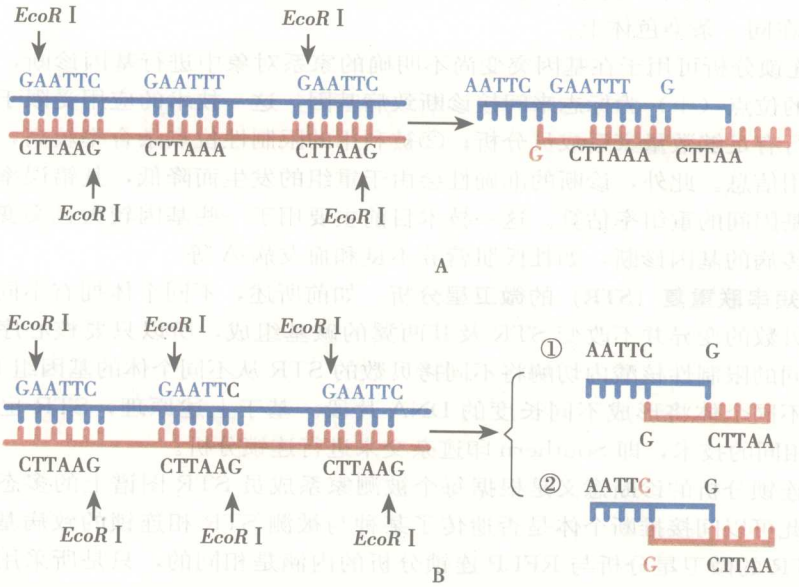


图 21-2　*EcoR* Ⅰ对 DNA 的酶切作用

A. 片段中有两个酶切位点，酶切形成单一的 DNA 长片段；B. 片段中部的单碱基突变（T→C）产生一个新 *EcoR* Ⅰ酶切位点，该片段出现三个酶切位点，酶切形成两个 DNA 短片段

（二）连锁分析的主要方法

1. RFLP 连锁分析　RFLP 连锁分析是一种建立在 RFLP 基础上的、与家系分析相结合的连锁分析技术。就其基本技术原理而言，就是利用某些遗传性疾病中疾病基因与同一条染色体上的某一 RFLP 之间的一种特殊关系，即连锁不平衡而进行的一种间接分析方法。在人群中，某一 RFLP 或者主要存在于具有某一疾病基因的染色体上，而很少存在于正常染色体上；或者主要存在于正常染色体上，而极少或根本不存在于具有这一疾病基因的染色体上，这种非随机的遗传现象称为连锁不平衡（linkage disequilibrium）。因此，利用家系中患病个体的染色体上，与致病基因紧密连锁的 RFLP 所导致的限制性位点的丢失或获得（图 21-2）为遗传标志，通过基于 RFLP 分析的 Southern 印迹（见本节相关内容），可以间接推断个体后代是否遗传了致病基因（图 21-3）。

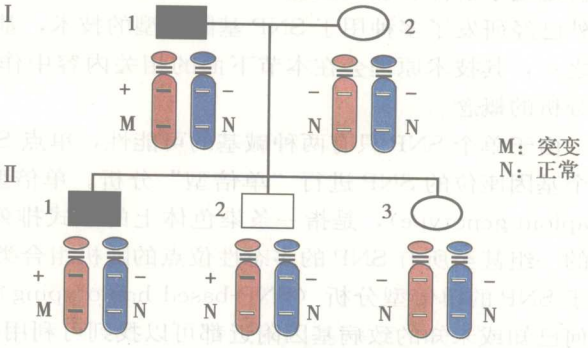


图 21-3　连锁分析的技术原理

在图 21-3 中，致病等位基因为显性基因（M），但不知道导致疾病

产生的基因突变。患病的父亲（图中的Ⅰ-1）为等位基因座位限制性位点的杂合子，有限制性位点标为（+），无限制性位点标为（-），位点（+）与突变基因（M）位于同一条染色体上。母亲为正常个体（Ⅰ-2），她的2条染色体上该座位的限制性位点均为（-），与正常等位基因（N）相连锁。该家系的长子（Ⅱ-1）获得了来自父亲的1条含突变基因和位点（+）的突变染色体和来自母亲的1条正常染色体，为患病个体；女儿（Ⅱ-3）则继承了携带位点（-）和N基因的2条正常染色体，为正常个体；次子（Ⅱ-2）虽然遗传得到了位点（+），但并未受累，这是因为发生了遗传重组事件，致使位点（+）和正常基因N出现在同一条染色体上。

RFLP连锁分析可用于在基因突变尚不明确的家系对象中进行基因诊断，即采用与突变基因连锁的位点（+）为标志来间接诊断致病基因，这一技术的应用受限于一些必要条件：①需进行有足够数量家系成员分析；②被利用的限制性位点杂合率较高，可以提供基因连锁的有用信息。此外，诊断的准确性会由于重组的发生而降低，其错误率可根据标志位点与疾病基因间的重组率估算。这一技术目前主要用于一些基因较大且突变类型不清楚的单基因遗传病的基因诊断，如杜氏肌营养不良和血友病A等。

2. 基于短串联重复（STR）的微卫星分析　如前所述，不同个体拥有不同重复次数的STR。但拷贝数的变异并不改变STR及其两翼的碱基组成，所以只要核心序列一致，就可以利用相同的限制性核酸内切酶将不同拷贝数的STR从不同个体的基因组DNA中切割下来，而且不同个体将形成不同长度的DNA片段。基于上述原理，STR也可以采用与RFLP作图相同的技术，即Southern印迹杂交来进行连锁分析。

同样，连锁分析的诊断意义是根据每个被测家系成员STR图谱上的多态性条带来进行判断，据此可以间接推断个体是否遗传了某种与被测STR相连锁的致病基因。不难理解，基于STR的微卫星分析与RFLP连锁分析的内涵是相同的，只是所采用的遗传标志物不同。为提高检测效率，STR分析目前已逐步被更为灵敏简便的PCR直接扩增技术-扩增片段长度多态性（amplified fragment length polymorphism，AFLP，Amp-FLP）分析法所取代。该方法只需要通过PCR技术，扩增致病基因内部或两侧与其紧密连锁的特定STR，经电泳检测扩增片段长度的多态性，即可快速实现基因诊断。由于STR分析在技术的简便性和分析结果所提供的信息上均优于RFLP分析，故该项技术已更广泛用于遗传病的连锁分析和基因定位。此外，基于STR的分析已成为法医学上个体识别的重要工具，又称为DNA指纹（DNA fingerprinting）技术（见第三节）。

3. 基于SNP的单倍型分析　如前所述，SNP的表现形式不是DNA片段的长度差异，而是以序列的变异为检测信息，因此SNP的分析技术有别于上述两类分析方法。但其实质还是检测基因序列的点突变，故用于点突变分析的技术也可用于SNP分析。由于SNP分析是基于群体和较大数量SNP位点的，故对技术的自动化提出了更高的要求，目前国内外已经研发了多种用于SNP基因分型的技术，基于DNA芯片的技术是重要的代表性技术之一，其技术原理会在本节下面的相关内容中作简要说明。这里，重点叙述SNP单倍型分析的概念。

由于单个SNP只有两种碱基的可能性，单点SNP所提供的信息有限，因此需要采用多个基因座位的SNP进行“单倍型”分析。单倍型（haplotype）一词来源于单倍基因型（haploid genotype），是指一条染色体上的顺式排列的等位基因座位。根据同一条染色体上的一组甚至所有SNP的多态性位点的随机组合类型来进行的分析技术（图21-4），称为基于SNP的单倍型分析（SNP-based haplotyping）。由于SNP数量巨大，分布广泛，在任何已知或未知的致病基因附近都可以找到可利用的SNP，以SNP的单倍型为遗传标志，可通过家系调查来进行致病基因的连锁分析或遗传制图，或可通过基于群体的SNP单倍

型分析，进行以寻找疾病易感基因为目的的相关性分析。当被测群体中某些 SNP 单倍型或某些 SNP 座位的组合的频率显著高于预期频率值时，即可判断为非随机相关的连锁不平衡事件，从而为寻找疾病易感基因提供有价值的遗传分析依据。由于采用多位点进行 SNP 分析可以避免因染色体互换、重排等偶然现象所造成的错误，因而会提高遗传分析和基因诊断的准确性。

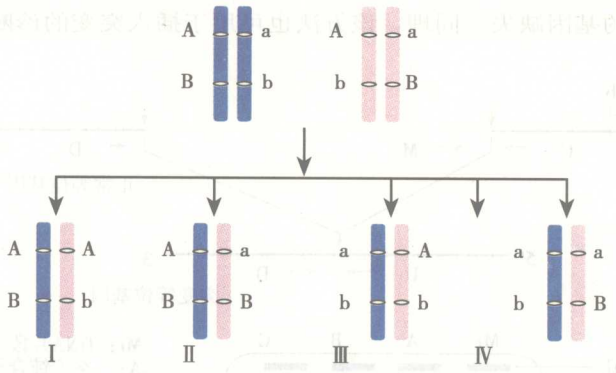


图 21-4　基于 SNP 的单倍型分析

两个 SNP 多态性位点构成的四种不同的单倍型（A-B、a-b、A-b 和 a-B）及其在一个家系中的分离情况。单倍型在家系中的传递符合孟德尔遗传规律，据此原理可以进行致病基因的连锁分析或遗传制图

二、针对遗传病致病突变的直接诊断

间接诊断是通过分析遗传标志的多态性来判断被检者患病的可能性，因而可能由于基因重组、家系成员不够完整、遗传标志杂合性所带信息量有限等因素而难以完成诊断。鉴于上述原因，针对致病基因及其基因突变谱已经清楚的遗传性疾病，可以采取直接检测致病基因突变的诊断策略。直接诊断的优点，一是不需依赖家系调查，在缺乏家系成员样品时也可实现对患病个体的诊断；二是检测方法更为简单可靠，因而结果更为直接和准确，直接诊断是目前用于临床遗传学诊断遗传病的主要方法学手段。以下介绍几种用直接诊断策略进行基因突变分析的一些代表性基因诊断技术：用于检测大片段的基因缺失或插入、点突变和 STR 的三种技术。

（一）基因缺失或插入的诊断

1. Southern 印迹法　用于检测大片段基因缺失或插入的经典方法是 Southern 印迹。该方法的基本原理和操作步骤在第十九章中已有介绍。在这一方法中，依据由放射自显影片段构成的 DNA 限制性酶切图谱（条带的大小和数量），可以区分正常和突变样品的基因型，并可获得基因缺失（或插入）片段大小等信息。Southern 印迹一般可以显示 50 bp～20 kbp 的 DNA 片段，片段大小的信息是该技术诊断基因缺陷的重要依据。

2. PCR 法　Southern 印迹实验结果可靠，但操作烦琐，费时费力，而且要使用放射性核素，这些因素使其难以作为一种常规的临床诊断手段得以广泛开展。因此，另一种与 PCR 技术相结合的方法，跨越断裂点的 PCR（又称裂口 PCR，gap-PCR），因其简便灵敏而被广为采用。该方法的思路是：设计并合成一组跨越突变（缺失或插入）断裂点的引物，将待测 DNA 样本进行 PCR 扩增，然后进行琼脂糖凝胶电泳，从扩增片段的大小直接判断是否存在缺失突变。

如图 21-5 所示，在缺失片段（如 20 kbp 缺失）的周围设计三个引物 U、M 和 D，共用引物 U 位于缺失基因 5′端断裂点的上游，引物 M 位于 5′端断裂点的下游，引物 D 位于 3′端断裂点的下游。这样，在正常基因中，引物 U 和 M 形成一对 PCR 引物，其扩增产物为 1 050 bp；引物 U 与 D 之间则因距离太远而无法形成 PCR 反应产物（5′和 3′端断裂点之间的距离为 20 kbp）；而在缺失基因中，由于 M 无结合位点而致 PCR 反应阴性，引物 U 与 D 因基因缺失而靠近，其扩增产物为 740 bp。通过上述三引物的双重 PCR 反应体系，即可直接诊断样品的基因缺失。同理，该方法也可用于插入突变的诊断。

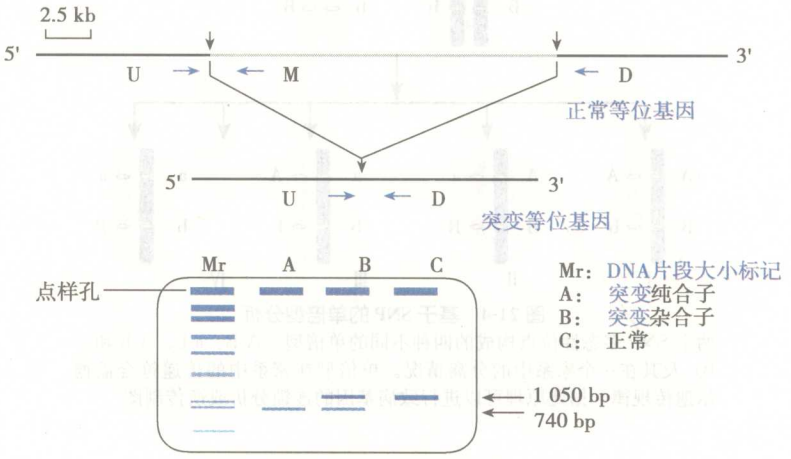


图 21-5　跨越断裂点的 PCR

向下黑色箭头示断裂点位置，两箭头间的灰色线条为缺失片段，U、M、D 为三个引物的位置；引物 U/M 扩增正常基因片段（1 050 bp），引物 U/D 扩增突变基因片段（740 bp）

（二）点突变的诊断

1. 等位基因特异性寡核苷酸分子杂交　检测点突变的有效技术是等位基因特异性寡核苷酸（allele specific oligonucleotide，ASO）分子杂交。使用该技术的前提条件是被测点突变的确切位置及其基因序列已经阐明，这样就可以设计合成包含突变位点在内的一对正常和突变的寡核苷酸探针。图 21-6 显示为 β 地中海贫血珠蛋白密码子 17 点突变检测的 ASO 探针序列。将这两个探针进行标记，再分别与待检 DNA 样品的扩增产物进行杂交和检测。正常人只能与正常序列的 ASO 探针杂交，突变纯合子只能与突变序列的 ASO 探针杂交，而突变杂合子则能与两种探针产生杂交信号。探针设计和杂交条件的摸索是建立 ASO 杂交技术的关键。

2. 反向点杂交　反向点杂交（reverse dot blot，RDB）是改进的 ASO 技术。ASO 分子杂交虽然可以准确地进行已知突变的基因分型，但由于一种突变需要对应一个探针和一组实验，对于突变类型较少的遗传病的诊断较为快速简便，当一种遗传病是由许多种点突变所引起，且其频率分散时，ASO 技术就显得很繁琐。为了能够同时检测多种突变，可以将针对各种突变和正常序列的 ASO 探针固定在杂交膜上，而将原来在 ASO 杂交体系中固定在膜上的 PCR 产物改为液相进行杂交，这种方法称为 RDB。利用这一方法，一次检测可以同时筛查多种突变，大大提高了基因诊断效率，已在一些常见遗传病，如 β 地中海贫血和囊性纤维化的基因诊断中得以应用。

近些年发展起来的 DNA 芯片（DNA chip）技术，其实质也是一种基于 RDB 的技术，但更为高效、高通量且操作易于自动化。DNA 芯片的基本原理是：将大量已知序列的

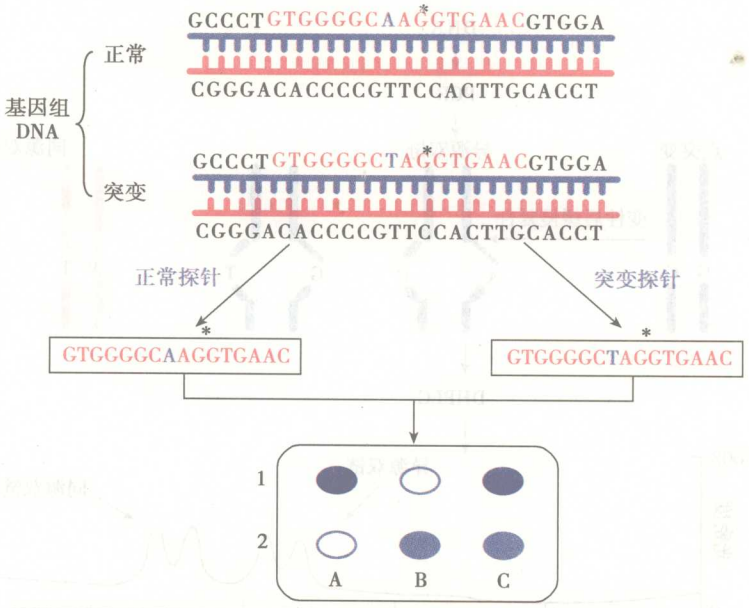


图 21-6 β地中海贫血 CD17 点突变的 ASO 分子杂交

1. 正常探针检测孔；2. 突变探针检测孔

每一纵行排列的两个点样斑代表一个样品，杂交时沿上下中线剪成两张膜条，分别与正常（上半部分）和突变（下半部分）探针杂交。三个样品的检测结果分别为正常（A）、突变纯合子（B）和突变杂合子（C）

DNA 探针依次排列并固定在硬质小基片（如玻片）上，在此固相载体上与荧光标记的扩增靶序列进行微杂交，杂交信号又经激光共聚焦扫描显微镜捕获，并通过计算机实现数据分析和报告检测结果。由于 DNA 芯片技术可以一次“集成”数量巨大的基因探针，故代表了基因诊断向快速、高效和自动化发展的趋势，是一种令人期待的有着广阔应用前景的基因诊断新技术。基于 ASO 点杂交的技术是我们理解诊断已知基因点突变的基础，也是点突变分析的经典技术。事实上，目前已经应用的点突变诊断技术已相当成熟和多样化，可根据需要选择相应的技术。

3. 变性高效液相色谱 在对临床病例进行基因诊断时，经常会遇到不能检测出已知类型突变的情况，如果表型明确指向某种疾病，可采用另一类筛查点突变的技术对目的基因进行基因序列扫描，以期发现和确定新的或未知突变类型。变性高效液相色谱（denaturing high performance liquid chromatography，DHPLC）是这类技术的代表之一。

DHPLC 技术的基本原理是利用 PCR 扩增过程的单链 DNA 产物可以随机与互补链相结合而形成双链的特性，依据最终产物中是否会出现异源双链来判断待测样品中是否存在点突变。如果被测 DNA 片段中不存在点突变，所有的 PCR 产物都将具有相同的序列，因此最终形成的所有的双链 DNA 都是一致的，即只产生一种同源双链。如果待测 DNA 片段存在着点突变，在 PCR 反应体系中就会产生 4 种不同的 DNA 双链分子，2 种为异源双链，另 2 种为同源双链。在给定的部分变性洗脱条件下，这些不同源的双链 DNA 片段的变性程度将有别于同源双链，由此造成的 DNA 分子电荷量等的差异可以在液相色谱柱中呈现出不同的滞留时间。因此，DHPLC 技术可将含不同的点突变的片段分离成不同的特征性洗脱峰而达到检测基因变异的目的（图 21-7），出现“变异”洗脱峰的样品可进一步通过 DNA 直接测序确定样品的突变位点和性质。这一技术依赖自动化操作的分析仪完成，目前已成为临床遗传学诊断的重要工具。

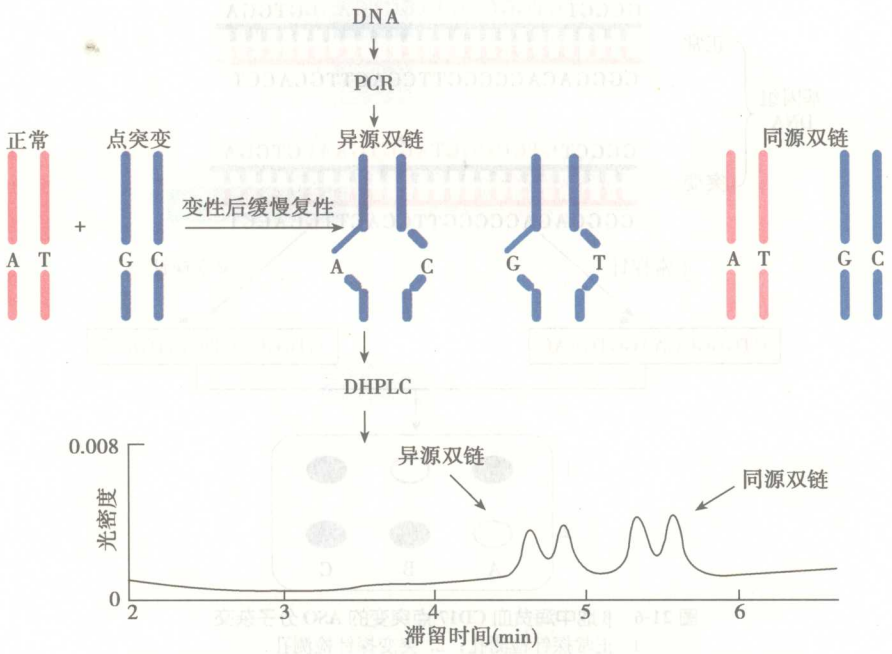


图 21-7　DHPLC 检测点突变的技术原理

（三）动态突变的诊断

短串联重复（STR）不仅是一种连锁分析的遗传标志，而且其拷贝数的增加也是一种导致某些遗传病发生的原因，称为动态突变（见第十章）。对于这类由 STR 扩增导致的遗传性疾病，可以采用 PCR 直接扩增技术来诊断基于 STR 的 DNA 片段长度多态性。该方法只需要以致病靶位点的 STR 及其周围序列为模板，设计合成一对针对 STR 两侧翼的寡核苷酸引物，样品经 PCR 扩增后，用聚丙烯酰胺凝胶电泳可直接检测扩增片段的长度，

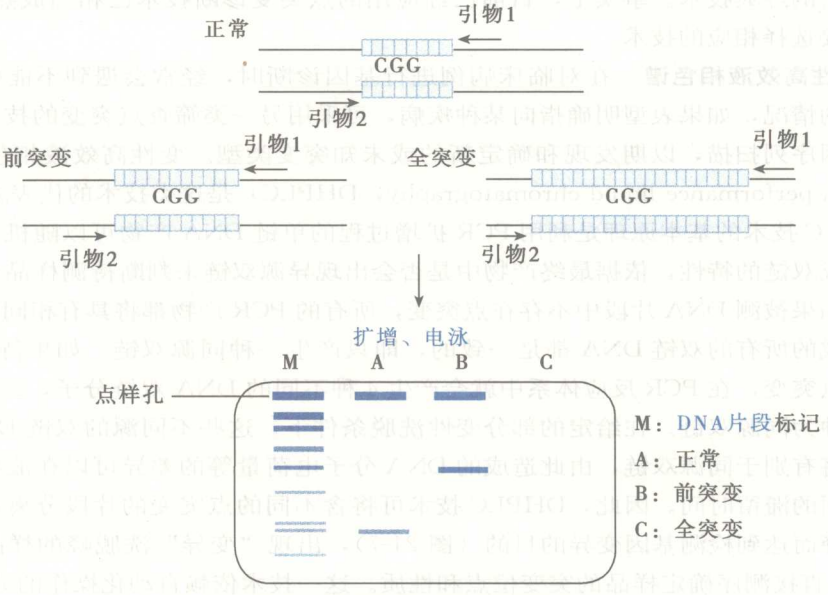


图 21-8　STR 突变序列的直接 PCR 诊断

据此计算出 STR 的串联重复拷贝数。例如脆性 X 染色体综合征，在基因的 5′端有一个 CGG 重复序列，正常人其拷贝数为 5～50，症状前携带者的拷贝数为 50～200，而脆性 X 染色体综合征患者的拷贝数超过 200。这些区别可通过一对针对 STR 两侧翼的寡核苷酸引物，经 PCR 扩增得到诊断（图 21-8），全突变因片段过长而无法扩增出条带。从这一技术的原理不难理解，STR 序列的 PCR 扩增也可用于基于 STR 差异的个体鉴定。

第三节 基因诊断的医学应用

当人类基因组的信息与人类疾病的关系完全明确以后，从理论上讲，基因诊断可以提供所有直接或间接涉及到基因结构/功能改变的疾病的诊断、预警和疗效预测。虽然目前距离这一目标仍有很长的路要走，但基因诊断在遗传病的临床和预防医学实践上已经获得了广泛的应用。以下主要阐述基因诊断在人类遗传病的诊断和预防方面的应用范围，其次简要介绍其在药物遗传学检测和法医学个体识别上的应用。

一、基因检测是疾病诊断和风险预测的有力工具

（一）遗传性疾病的诊断性检测和症状前检测

在欧美发达国家，遗传病的基因诊断，尤其是单基因遗传病和某些恶性肿瘤等的诊断，已成为医疗机构的常规项目，并已形成在严格质量管理系统下的商业化服务网络。目前已列入美国华盛顿大学儿童医院和区域医学中心主持的著名基因诊断机构-GENETests 网站（http：//www. geneclinics. org/）检测名单的人类遗传病基因服务项目多达 1 170 种。针对单基因病（包括少数肿瘤）的诊断性检测（diagnostic testing）是遗传病基因诊断的主要应用领域，对于这类检测，基因诊断可以提供临床确诊的信息。染色体病是一类常见的遗传性疾病，长期以来，其诊断依据除了临床表现，主要靠细胞学核型分析和生化检查结果。由于细胞学技术的操作相对繁杂且耗时，而生化检查结果又缺乏特异性，采用分子诊断技术进行某些染色体病的诊断性检测已成为一项十分有用的可选择方案。此外，还可针对一些特定疾病的高风险个体、家庭或潜在风险人群进行症状前检测（pre-symptomatic testing），其目标为预测个体某种（些）疾病的未来发生状况。这种诊断是随着对遗传病分子机制认识水平的提高和分子诊断技术的介入而发展起来的，它所针对的疾病为一类迟发性（delayed age onset or late onset）遗传性疾病，如面肩肱型肌营养不良。同时，这一检测也可为某些可治疗的遗传病提供有价值的早期、症状前诊断结果，为挽救患者争取宝贵的时间。如迟发性显性遗传性感音神经性耳聋和 X-连锁迟发性脊椎骨骺发育不良等。

我国基因诊断的研究和应用始于 20 世纪 80 年代中期，目前主要开展针对一些常见单基因遗传病的诊断性检测，如地中海贫血、甲型血友病、进行性肌营养不良症等，表 21-1 列举了在我国开展的一些代表性常见单基因病基因诊断及其方法学案例。

（二）多基因常见病的预测性诊断

对于一些多基因常见病，通过基于 DNA 分析的预测性诊断可为被测者提供某些疾病发生风险的评估意见，如由于 *BRCA1* 和 *BRCA2* 基因突变可提高个体的乳腺癌发病风险，其基因诊断已成为一些西方国家中人群健康监测的项目之一。随着基因变异和疾病发生相关研究的知识积累，针对肿瘤和其他一些多基因常见病的这类预警性风险预测诊断正在逐步走入临床。在一些有明显遗传倾向的肿瘤中，肿瘤抑制基因和癌基因的突变分析，是基因检测的重要靶点。与肿瘤发生密切相关的这两类基因如：①视网膜母细胞瘤与 *RB* 基因；

表 21-1　我国一些代表性常见单基因遗传病分子诊断

疾病	致病基因	突变类型	诊断方法
α地中海贫血	α珠蛋白	缺失为主	Gap-PCR、Southern 杂交、DHPLC
β地中海贫血	β珠蛋白	点突变为主	反向点杂交、DHPLC
甲型血友病	凝血因子Ⅷ	点突变为主	PCR-RFLP、SSCP
乙型血友病	凝血因子Ⅸ	点突变、缺失等	PCR-STR 连锁分析
葡萄糖-6-磷酸脱氢酶缺陷症	*G6PD* 基因	点突变	SSCP、DHPLC
苯丙酮尿症	苯丙氨酸羟化酶	点突变	PCR-STR 连锁分析、ASO 分子杂交、SSCP
Marfan 综合征	原纤蛋白基因	点突变、缺失	PCR-VNTR 连锁分析、DHPLC
DMD/BMD	*DMD/BMD*	缺失、点突变、STR 重复	多重 PCR、PCR-STR 连锁分析、DHPLC、MLPA
肝豆状核变性	*WD*	点突变、缺失、插入	PCR-RFLP、PCR-酶切分析、DHPLC
常染色体显性遗传性多囊肾病	*PKD1*、*PKD2* 等	点突变	PCR-RFLP、PCR-STR 连锁分析、DHPLC

②肾母细胞瘤与 *WT1* 基因；③家族性结肠癌与 *APC* 基因；④Li-Fraumeni 综合征与 *P53* 基因；⑤乳腺癌与 *BRCA1* 和 *BRCA2* 基因。研究表明 2 型糖尿病和高血压等常见病均有多个参与疾病发生和发展的相关基因，分析这些基因的多态性或基因突变，可为疾病易感性和发病风险提供评估指标，对于有家族史的个体而言，预测性基因诊断结果是开展临床遗传咨询最重要的依据。相信在相关基础研究取得进展的基础上，多基因常见病的预测性诊断会越来越多地得到应用，成为未来医学中疾病诊断的主要内容之一。

（三）传染病病原体检测

病原体种类繁多，包括病毒、细菌、真菌、支原体、衣原体、立克次体、螺旋体以及寄生虫等。病原体的基因诊断是针对病原体自身特异性核酸（DNA 或 RNA）序列的分子诊断。基因诊断的目的就是通过分子杂交和基因扩增等手段，去鉴定和发现这些独特的外源性基因组、基因或基因片段在人体组织中的存在，从而证实某种病原体的感染与否。应用于微生物基因检测的技术选择原则是：①需复杂营养培养的微生物的鉴定；②以流行病学等为目的分离株的快速分型；③微生物耐药性的快速检测；④不可培养的微生物的鉴定。

基因诊断已经成为现代医学中病原学诊断的重要组成部分，它的出现是对临床医学中传统的病原学检测的一项有价值的补充。传统的病原学检测包括形态学观察、免疫学检测和培养等方法。它们都有各自的技术优势和应用范围，但同时又有局限性，如采用光学显微镜的形态学观察，难以满足大量标本检测的要求；病原体的分离培养需要较长时间；采用免疫学方法检测抗体不能用于早期诊断等。分子诊断技术的特点决定了病原体的基因诊断较传统方法有更高的特异性和敏感性，有利于疾病的早期诊治、隔离和人群预防。但由于基因诊断只能判断病原体的有无和拷贝数的多少，难以检测病原体进入体内后机体的反应及其结果，因此，基因检测并不能完全取代传统的检测方法，它将与免疫学检测等传统技术互补而共存。

二、基因诊断是遗传病预防的重要技术手段

（一）遗传筛查和产前诊断

遗传筛查是指针对特定人群以发现某些遗传病和先天性缺陷（congenital defects）的高风险个体为目标的系统检查，对于常染色体和X-连锁的隐性遗传病而言，遗传筛查的目的是发现杂合子个体，也可称为携带者检测（carrier testing）。通过遗传筛查检测出的高风险夫妇（或拟婚对象）需给予遗传咨询和婚育指导，在“知情同意”的原则下于适宜的妊娠期开展胎儿的产前诊断（prenatal diagnosis），若胎儿为某种严重遗传病的受累者，可在遗传咨询的基础上由受试者决定“选择”终止妊娠，从而在人群水平实现遗传性疾病预防的目标。由于目前人类对遗传病和先天性缺陷还没有理想的治疗办法，采用产前诊断技术预防遗传病患儿的出生，是国际社会公认的降低遗传病发生率的主要途径。而产前诊断的主要技术基础是基于DNA分析的基因检测。目前遗传筛查所用的技术一般为表型分析方法，随着分子诊断技术的发展和不断完善，其快速简便和自动化的特点为人群筛查提供了新的技术选择，目前已有针对若干种遗传病的“分子筛查”技术应用于人群遗传病的预防实践中。如静止型α地中海贫血和脊肌萎缩症的携带者，采用表型指标不能准确诊断出杂合子，用基因诊断技术可直接检测致病基因的缺陷所在从而实现筛查的目标。与表型筛查相比，分子筛查更加直接、准确，是未来遗传筛查的方向。采用DNA基因分型技术，对采自羊水、绒毛或脐带血等来源的DNA进行遗传病的诊断性检测，目前已经可对大多数重要单基因遗传病开展产前诊断，如地中海贫血、囊性纤维化、镰状贫血症、苯丙酮尿症、血友病、进行性肌营养不良症和脆性X综合征等，都已成为产前诊断的对象。以选择性淘汰受累胎儿为目的的产前基因诊断已逐步被广大公众所接受，并且会越来越多地应用于各种人类遗传病的诊断，对人类生殖健康观念的进步产生了深刻的影响。

（二）植入前遗传诊断

植入前遗传学诊断（preimplantation genetic diagnosis，PGD）是辅助生殖技术与遗传学诊断技术相结合而形成的一种孕前诊断技术，它针对经体外受精（in vitro fertilization，IVF）获得的胚胎在移入宫腔之前进行快速的遗传学分析，筛选出健康胚胎个体，以期达到避免某些单基因遗传病和先天性缺陷异常妊娠的目的。与传统的产前诊断相比，这一技术最大的优势是主动筛选出健康胚胎，而非在产前诊断后被动取舍胎儿。基因诊断是该项技术的核心内容，即应用基于PCR或FISH等快速、敏感和准确的分子诊断技术，在受精卵分裂到4～8个细胞期时，对用显微操作仪吸取的1～2个卵裂球细胞实施快速的遗传病学诊断（通常需在6～8小时内完成诊断），从而完成对所检胚胎的鉴定，而这一过程不影响胚胎的继续发育。PGD是人类预防先天性出生缺陷的主要对策和技术手段之一，自1989年PGD应用于临床以来，该技术目前已在预防数十种人类遗传性疾病的临床实践中取得成功。其中，FISH技术主要应用于几种三体综合征（如21三体、11三体、13三体、18三体）和性染色体数目异常（如Turner综合征、Klinefelter综合征、XYY综合征）等染色体病的PGD筛查上。PCR技术则可用于大多数单基因病（包括常染色体显性及隐性遗传病、少数性连锁单基因遗传病）的植入前基因诊断，如囊性纤维化、黑矇性白痴、进行性肌营养不良、新生儿溶血、甲型血友病、马方综合征、β地中海贫血等，并由此衍生出建立在分子杂交和PCR基础上的多种分子诊断技术。PGD技术目前主要用于染色体病和为数不多的人类常见单基因病的诊断和干预，相信PGD技术的进一步发展将对推动人类生殖健康，预防多种遗传性疾病的发生发挥越来越大的作用。

三、基因诊断可用于疗效评价和用药指导

遗传诊断还可应用于临床药物疗效的评价及提供指导用药的信息。例如，急性淋巴细胞性白血病经化疗等综合治疗后，大部分患者可获得缓解，但容易复发。白血病复发的主要根源在于患者体内少数残留的白血病细胞（约$<0.05\times10^9$/L）。临床上目前采用的方法中，包括流式细胞仪分析、FISH 检测和 PCR 直接扩增，相比较而言，染色体分析和分子诊断技术优势明显，FISH 是公认的可靠性最好的方法；PCR 技术理论上可以检测出单拷贝的基因模板，敏感性最高且简单快速，已成为临床上检测和跟踪微小残留病（minimal rudimental disease，MRD）的常规方法，其分析结果是临床医生预测白血病的复发、判断化疗效果和制订治疗方案很有价值的参考指标。

人群中对药物的反应性存在着个体差异，致使药物的不良反应难以避免。研究证实，氨基糖苷类抗生素致聋的病因学与线粒体 DNA *12S rRNA* 基因第 1555 位 A→G 同质性点突变有关，在人群中通过分子筛查发现这种突变的个体，对指导医生避免使用氨基糖苷类抗生素，防止儿童产生药物中毒性耳聋有很好的参考价值。造成药物作用个体差异的原因之一是由于药物代谢酶类（如细胞色素 P450）的遗传多态性，也可以理解为个体对某些药物的反应性差异取决于其所具有的先天性基因多态性或其单倍型的“遗传素质”，因此，通过测定人体的这些基因多态性或其单倍型可以帮助我们预测药物代谢情况或疗效的反应性，从而制订针对不同个体的药物治疗方案，提高疗效，减少药物不良反应。目前，对药物代谢酶及其他相关的药物治疗靶点的基因多态性正在进行广泛深入的研究，在系统阐明人类药物代谢酶类及其他相关蛋白的编码基因遗传多态性的基础上，我们可以通过对不同药物代谢基因靶点的药物遗传学检测（pharmacogenetics testing），为将来在临床上真正实现药物的剂量个体化治疗提供技术支撑。

四、DNA 指纹鉴定是法医学个体识别的核心技术

基因诊断在法医学上的应用，主要是采用基于 STR 的 DNA 指纹技术进行个体认定，该方法已成为刑侦样品的鉴定、排查犯罪嫌疑人、亲子鉴定和确定个体间亲缘关系的重要技术手段。DNA 指纹的遗传学基础就是 DNA 的多态性，世界上除了部分同卵双生子外，人与人之间的某些 DNA 序列特征彼此互不相同，具有高度的个体特异性和终生稳定性，正如人的指纹一般，故称为 DNA 指纹。

DNA 指纹技术建立于 1985 年，英国遗传学家 A. Jeffreys 采用小卫星 DNA（即 20～

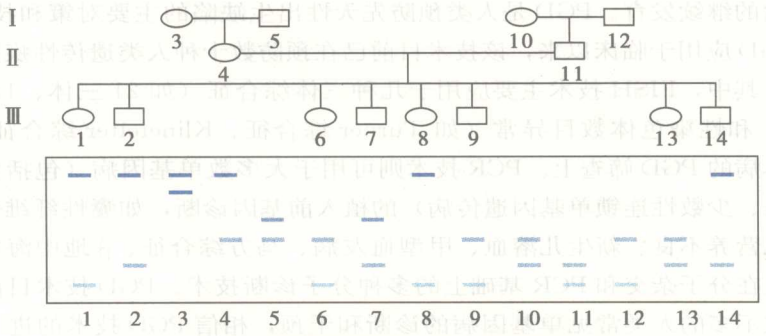


图 21-9　STR 等位基因片段在一个人类家系中的分离

70 bp 的寡核苷酸串联重复序列）作探针进行人基因组 DNA 的 Southern 印迹分析，在一个家系中获得了一张具有高度多态性的基于 STR 的 DNA 图谱，图谱中的每一个成员都有各自独特的 DNA 电泳谱带。在该家系图谱中，我们可以清楚地看到各子代成员的区带都是由亲代遗传而来，半数与父亲的区带一致，半数与母亲的区带一致，完全符合孟德尔遗传规律（图 21-9）。当前，基于 PCR 扩增的 DNA 指纹技术已取代了上述基于 Southern 杂交的操作程序。选择若干个基因位点（如 STR、HLA 位点等）设计相应的 PCR 引物对，对待测 DNA 样本进行 PCR 扩增和带型比较后即可判断结果。该方法快速、灵敏，可以对微量血痕、精液、唾液和毛发进行个体鉴定。

（徐湘民　商　璇）

第二十二章　基因工程药物与疫苗

关键知识点

- 基因工程：又称 DNA 重组技术或分子克隆技术，是指将一种生物体（供体）的基因与载体在体外进行拼接重组，然后转入另一种生物体（受体）内，使之按照人们的意愿稳定遗传并表达出新产物或新性状的 DNA 体外操作程序。
- 基因工程药物/疫苗：是指通过基因工程方式制备得到的具有疾病治疗或预防价值的重组蛋白质/多肽，DNA 或 RNA。
- 基因药物：又称基因转录表达调控药物，包括两方面的内容。一是指将具有治疗意义的 DNA 重组入真核表达载体，直接转移至人体细胞，在体内表达出具有治疗作用的蛋白质/多肽；二是应用直接作用于细胞内 DNA 或 RNA、抑制基因复制和表达的核酸片段或其人工合成的类似物，以及对 RNA 具有特异切割作用的核酶等。

关键词： 基因工程；基因工程药物、基因工程疫苗、基因药物

20 世纪 70 年代以来，以 DNA 重组为核心的现代生物技术-基因工程技术，为生命科学的发展带来了革命性的变化，并极大地推动了生物制药产业的发展。传统的生物制药主要是指利用动物、植物、微生物等生物材料加工制成的含有某些或一组天然活性成分的复合制剂，或利用生物化学和药理学原理，应用近代分离、纯化技术从生物体制取的具有针对性治疗作用的特异性单一活性成分。但生物材料的来源、成本以及免疫原性等问题使生物药物和疫苗的制备和应用受到了限制。以基因工程为代表的第三代生物制药技术则克服了上述不足。基因工程产品是指运用现代生物技术，从动物（包括人体）、植物和微生物中获得有实际应用价值的基因（如药源基因），或通过基因操作技术，得到具有更高药理活性的天然结构类似物或全新结构的基因（包括 DNA 和 RNA），然后再将上述基因在体外培养的动物、植物和微生物细胞中，或直接在动物、植物体内进行大量表达而制备得到的生物活性更高、免疫原性降低、毒副作用更低的重组蛋白和具有治疗作用的 DNA/RNA。

第一节　基因工程药物和疫苗的种类及制备体系

一、基因工程药物与疫苗的种类

按性质划分，广义的基因工程药物包括蛋白质/多肽类药物和疫苗、核酸类药物和疫苗。前者主要包括基因工程（重组）细胞因子类、激素类、抗菌肽类、酶和酶抑制剂类、治疗性抗体类药物以及基因工程微生物、寄生虫和肿瘤疫苗，后者则包括裸 DNA 基因药物、反义核酸类药物和核酸疫苗等。

二、基因工程药物和疫苗的制备体系

基因工程药物和疫苗制备的关键是目的蛋白或 RNA 表达系统的选择。表达系统包含

运载外源基因的载体和宿主细胞两部分。通过将目的基因克隆至合适的载体，并根据产物的特点选择合适的宿主表达细胞，可以制备相应的基因工程药物与疫苗。宿主细胞应满足以下条件：①宿主细胞及其载体的遗传背景清楚，便于进行DNA重组操作；②宿主细胞的代谢途径及其调控机制清楚，容易进行代谢调控；③容易获得较高密度的细胞培养；④能利用廉价而容易获得的原料；⑤产热少、耗氧低，发酵温度适当；⑥产量和得率高，产物容易提取纯化。用于基因表达的宿主细胞包括原核细胞（大肠杆菌，枯草芽孢杆菌和链霉菌等）和真核细胞（酵母，丝状真菌，昆虫细胞和哺乳动物细胞等）两大类。常用的表达系统及其特点如表22-1所示。

表22-1　常用的基因工程药物和疫苗制备体系

宿主细胞	常用载体	特点
大肠杆菌	质粒：pMB1（pUC）类、p15A类、ColE类、pSC101类	优点：目的蛋白表达水平高、培养周期短；可进行融合表达 缺点：缺少蛋白质翻译后修饰和加工；表达产物多以不溶性的包含体形式存在；细胞壁中含内毒素成分
枯草杆菌	质粒：pUB110、pC194、pE194 噬菌体：Φ105、SPβ	优点：目的蛋白表达水平高、培养周期短；表达产物可外分泌；不含内毒素成分 缺点：分泌大量蛋白酶，可使外源表达产物发生降解；质粒不稳定
链霉菌	质粒：pIJ101、pIJ702、pIJ922、pIJ940、pHJL197、pHJL210（SCP2*/pBR322）、pHJL302（SCP2*/pUC19）、pIH1351 柯斯载体（cosmid） 噬菌体：KC系列	优点：目的蛋白表达水平高、培养周期短；表达蛋白大量分泌；不含内毒素成分；自身分泌的蛋白酶低 缺点：质粒不稳定
酵母	穿梭质粒：整合型质粒（YIp），附加体质粒（YEp），复制型质粒（YRp），着丝粒质粒（YCp）	优点：表达产物可进行一定的翻译后修饰与加工；易培养 缺点：糖基化与真核细胞显著不同，含有过多的甘露糖或糖链过长
哺乳动物细胞	质粒载体：pcDNA3.1，pSI，pCMV-HA，pBudCE4.1，pTRE 病毒载体：腺病毒、腺相关病毒、牛痘病毒和杆状病毒等	优点：表达产物可进行翻译后修饰与加工 缺点：难以大规模培养
昆虫细胞	质粒载体：pDS47、pAc5.1、pMT、pIB、pIZt、pIZ、pMIB；pIE1-1、pIE1-2、pIE1-3、pIE1-4pBAC-5、pBACgus-5、pAcP（+）IE1-1、pAcP(−)IE1-5 杆状病毒载体：BAC-To-BAC®系统	优点：表达产物可进行翻译后修饰与加工 缺点：难以大规模培养

第二节　基因工程蛋白质/多肽药物

一、细胞因子类药物

细胞因子是一类由免疫活性细胞及相关细胞分泌的具有调节细胞功能的高活性多功能小分子蛋白质。通过与细胞表面的特异性受体相互作用后发挥广泛的生物学作用包括杀伤

肿瘤细胞或诱导其凋亡、抗病毒作用、刺激骨髓的造血功能以及机体的免疫应答等，其中的一些因子还参与机体的炎症反应。体内多种细胞因子间的相互作用组成复杂的细胞因子网络，从而发挥机体对内、外环境变化的应答和调节。

由于细胞因子在体内的含量很低，纯化困难，所以虽然在20世纪50年代末就发现了第一个细胞因子——干扰素，但直到20世纪70年代末随着分子生物学技术的发展才使细胞因子的研究和应用得以飞速发展。利用基因工程技术，许多新的细胞因子被发现和大量表达，在短短数年之内已有数十个细胞因子被批准上市，用于肿瘤、感染和造血功能障碍等疾病的治疗。

细胞因子包括干扰素（interferon，IFN）、白细胞介素（interleukin，IL）、集落刺激因子（colony-stimulating factor，CSF）、肿瘤坏死因子（tumor necrosis factor，TNF）、转化生长因子（transforming growth factor，TGF）、趋化因子（chemokine）等6个家族。目前上市的细胞因子类药物及其适应证如表22-2所示。

表22-2　主要的治疗用细胞因子类药物及其适应证

细胞因子	亚类	适应证
IFN	IFN-α	毛细胞白血病，慢性病毒性感染
	IFN-β	多发性硬化症
	IFN-γ	类风湿性关节炎
IL	IL-2	某些肿瘤，感染
	IL-11	血小板减少症
CSF	G-CSF	中性粒细胞减少症
	GM-CSF	造血功能障碍，中性粒细胞减少症
	Multi-CSF	骨髓衰竭，再生障碍性贫血
	EPO	慢性肾功能衰竭引起的红细胞减少症
	TPO	放疗、化疗引起的血小板减少症
	SCF	造血功能障碍
TNF	TNF突变体	某些肿瘤
GF	EGF（外用）	角膜溃疡，烧、创伤愈合
	bFGF（外用）	组织修复，伤口愈合
	PDGF（外用）	创伤、糖尿病性或褥疮性溃疡愈合
	IGF-1	2型糖尿病

二、激素类药物

激素是由内分泌细胞分泌的高效能活性物质。在体内的含量很少，经血液循环到靶组织后作为一种化学信使或信号分子发挥专一的生理效应。按照化学结构激素可以分为蛋白质/多肽类激素和类固醇激素。重组激素目前批准上市的有重组人生长激素、重组人胰岛素、重组人促卵泡激素和重组人甲状旁腺激素。

（1）生长激素（GH）：GH是由垂体前叶分泌的含两对链内二硫键的、由191个氨基酸残基组成的单体蛋白质。GH在体内有多种极重要的功能如促进物质代谢、刺激新骨形成、刺激红细胞生成、增强免疫系统功能、减少脂肪及其在体内的分布等。临床上自1958年以来一直使用尸体或动物来源的GH治疗侏儒症。1985年FDA批准由大肠杆菌表达的第一代基因工程GH上市；第2代重组人GH是分泌表达型产品，可产生与天然蛋白完全一致的一级结构和高级结构，生物学活性因而提高。目前重组人GH已国产化。

（2）胰岛素（insulin）：胰岛素是以前胰岛素原的形式由胰岛B细胞分泌的，分泌后

形成前导序列（信号肽，23 个氨基酸）和 A 链（21 个氨基酸）、B 链（30 个氨基酸）、C 肽（25～38 个氨基酸），经跨膜运输，最后在高尔基体内形成二硫键，并形成二聚体、三聚体或四聚体，进而分泌入血液中发挥调节代谢的作用。临床上胰岛素主要用于治疗胰岛素依赖型（Ⅰ型）糖尿病。重组人胰岛素早在 1982 年即被批准上市，当时采用的技术是分别在大肠杆菌中表达 A、B 链，然后在体外化学连接为有活性的胰岛素，现在则是在大肠杆菌直接表达人胰岛素原，体外再转化为成熟的胰岛素。

（3）促卵泡激素（follicle stimulating hormone，FSH）：体内 FSH 和黄体生成素（luteinizing hormone，LH）合称为促性腺激素。FSH 是分子量为 24 kDa 蛋白，其主要作用是促进精子及卵子的生成，促进卵巢发育，故临床上可用于下丘脑-垂体功能障碍和多囊卵巢综合征。研究表明，FSH 在垂体 LH 协同下，能促进卵巢卵泡生长发育和雌激素分泌，引起正常发情，提高受胎率。美国已批准重组人 FSH 上市，我国亦已批准进口。国产 FSH 已进入临床试验。

（4）甲状旁腺激素（parathyroid hormone，PTH）：PTH 是由甲状旁腺主细胞分泌，由 84 个氨基酸残基组成。PTH 的主要生理功能为促进骨质溶解，动员骨钙入血，血钙增高，骨和血中碱性磷酸酶活力增加；抑制肾小管对磷的再吸收，促进尿磷排出增多，血磷降低；PTH 通过活化维生素 D_3 间接促进肠粘膜吸收钙、镁以及磷。Eli Lilly 公司生产的重组人 PTH 衍生物（商品名 FORTEO）含 34 个氨基酸，与人 PTH 的 N 端序列相同，具有调节成骨和破骨细胞作用的平衡、促进骨质重塑的作用，临床上用于治疗和预防妇女绝经后或男性性腺障碍引起的骨质疏松症。

三、基因工程抗体药物

治疗性单克隆抗体技术是当前国际生物技术领域开发的热点，它被视为新药研究和开发的又一个重要方向。

（一）抗体药物研制的发展

人体的免疫系统具有特别的敏感性和专一性。长期以来人们一直在探索将它的这些特性应用于疾病的治疗。1975 年，G. Kohler 和 C. Milstein 发现单克隆抗体（monoclonal antibody，mAb）可与肿瘤抗原发生特异和专一性的结合后，人们才利用这一技术生产具有特定药用价值的抗体。近 30 年来，科学家们一直尝试利用 mAb 技术生产特异性强的药物来治疗肿瘤、感染性及自身免疫性等疾病。1986 年 FDA 批准了第一个用于抗异体移植排斥反应的 mAb-Orthoclone，由于抗体制造技术的不断突破，自 1994 年来已连续批准 20 余个治疗性 mAb。

当前特别重视肿瘤治疗性 mAb 的研究。抗体和肿瘤细胞结合后，有的会导致肿瘤细胞被免疫细胞识别而被杀伤、吞噬；有的还可以直接抑制细胞的增殖。例如针对癌基因 HER2 的单克隆抗体贺赛汀（Herceptin）就具有抑制 HER2 介导的生长因子信号，促进 HER2 的降解等作用。双特异性 mAb 可以专一性的识别一个与癌症相关的抗原，同时，它的另一个特异性识别位点可以结合一个引发细胞毒性的分子，如白细胞表面受体或化疗药物等。应用这类 mAb 后，免疫效应细胞如 T 淋巴细胞或巨噬细胞与化疗药物同时可以通过抗体携带到靶细胞上而定向发挥功能。

多种人肿瘤细胞表面都有相对于正常细胞过量表达的生长因子受体如 EGFR、HER-2 等。因此，利用这些受体可开发成能特异性定位的双功能抗体并进行肿瘤的治疗。如可特异性识别 EGF 受体并同时具有激活 CD3 功能或与阿霉素结合的双功能抗体，特异性识别 HER-2 并可激活 CD3、CD64（FcγRⅠ）、CD32（FcγRⅡ）、CD16（FcγRⅢ）等的双功

能抗体等。

（二）基因工程抗体

基因工程或重组抗体技术就是在基因水平对抗体分子进行切割、拼接或修饰，甚至将人工全合成的抗体基因导入受体细胞表达，产生新型抗体。

基因工程抗体包括鼠 mAb 的人源化（humanized）、小分子抗体（Fab 抗体、单链抗体、单域抗体和超变区多肽）、特殊类型的重组抗体（双价抗体、免疫粘连素、催化抗体）等。

（三）人源化抗体

1. 噬菌体抗体库　随着 PCR 技术的发展，利用基因工程的方法可将全套人抗体重链和轻链 V 区基因克隆出来，以噬菌体为载体，将抗体基因（Fab 或 ScFv 基因等）与噬菌体编码的外壳蛋白Ⅲ或Ⅷ相连，在噬菌体表面以抗体-外壳蛋白融合形式表达。用不同的抗原可筛选出与之特异性结合的相应抗体。

2. Xeno mouse 产生的人源化抗体　噬菌体抗体库制备的抗体通常为 Fab 片段或 ScFv 抗体，由于没有 Fc 段，亲和力受到极大影响。Xeno mouse 制备人源化抗体的原理是：将小鼠的全套抗体基因去除，同时将人的轻链、重链基因插入到鼠抗体基因相应的染色体部位，这样用抗原刺激小鼠时就发生人抗体基因的重排，从而产生人源化抗体。

（四）全人单克隆抗体——人-人杂交瘤技术

单克隆抗体技术的研究经过了 4 个阶段：鼠源性抗体、嵌合抗体、人源化抗体及全人抗体。治疗性抗体药物最主要的副作用来源于抗体内的非人体成分，主要是鼠源性蛋白。这种蛋白进入人体后成为抗原，导致人体产生免疫反应，出现多种副作用。理想的全人治疗性抗体应不含鼠源性蛋白成分，因此单纯从技术上来说，全人抗体技术是最先进的治疗性抗体开发技术。

制备全人 mAb 的平台为人-人杂交瘤技术，这项技术的关键就是将人 B 细胞与人骨髓瘤细胞株（Karpas 707）进行融合，经培养和筛选后获得产生全人 mAb 的杂交瘤细胞株。人-人杂交瘤抗体技术与其他抗体技术相比具有以下优点：①没有鼠源性异种蛋白，副作用更小；②无需抗体分子结构的变构转换及拼接等，技术更为简便，高通量筛选，开发周期短；③生产工艺简单，抗体产量很高，达 200～700 mg/L；④与人体内天然产生的抗体完全一致，具有经过自然筛选的合适的亲和力（affinity）和结合力（avidity）。

四、溶栓和抗凝血药物

溶栓药物的作用机制是通过激活血浆纤溶酶原，形成有活性的纤溶酶，后者催化血栓主要基质纤维蛋白发生水解，从而使血栓溶解。临床上可用于急性心肌梗死和急性脑血栓的抢救。目前国内外已正式批准的主要溶栓药物有：提取的链激酶（streptokinase，SK）和尿激酶（urokinase-type plasminogen activator，u-PA）；化学修饰的对甲氧苯甲酰纤溶酶原链激酶激活剂复合物（anisoylated plasminogen streptokinase activator complex，APSAC）；重组的组织型纤溶酶原激活剂（recombinant tissue-type plasminogen activator，rt-PA）、链激酶（recombinant streptokinase，r-SK）和葡激酶（recombinant staphylokinase，r-SAK）。

1. 组织型纤溶酶原激活剂　组织型纤溶酶原激活剂（t-PA）是由 527 个氨基酸残基组成的糖蛋白，含 17 对二硫键和 3 个糖基化位点。天然 t-PA 为单链分子，受纤溶酶或胰蛋白酶催化水解 Arg275-Ile276 的肽键，转变为由二硫键相连的双链 t-PA。其 N 端称为 A

链，C端称为B链，后者含t-PA活性中心，即His322-Asp371-Ser478。糖基化位点出现在Asn117、184和448。t-PA与血栓中的纤维蛋白有极高的亲和力（Km＝0.11 μmol/L），与纤维蛋白结合后的t-PA对纤溶酶原的亲和力显著提高（Km＝0.05～0.16 μmol/L，Kcat＝0.1～0.35 s^{-1}），t-PA-纤维蛋白-纤溶酶原在血栓部位形成三元复合物，有利于激活纤溶酶原，形成纤溶酶，水解复合物中的纤维蛋白，溶解血栓。由于t-PA分子内含有17对二硫键，复性困难，故重组t-PA常采用哺乳动物细胞表达。

2. 链激酶　链激酶（SK）是由溶血性链球菌产生的一种含414个氨基酸残基的单链多肽。SK是非蛋白酶类纤溶酶原激活剂，当SK与纤溶酶原结合后引起构象发生改变，使纤溶酶原活性位点暴露而受到活化成为纤溶酶，结果形成的SK-纤溶酶复合物起到溶血栓的作用。重组SK已获准生产，用于临床救治急性心肌梗死患者。尽管SK溶栓活性高，但对人体具有抗原性。SK结构与功能关系的研究发现，其C端的42个氨基酸（373～414）与抗原性有关，通过蛋白质工程技术，可获得C端缺失的重组SK衍生物。国内通过基因改造和重组表达出比活性高而副作用低的重组SK，并于1998年获准生产。

3. 葡激酶　葡激酶（SAK）是由金黄色葡萄球菌受溶源性噬菌体感染后合成和分泌的一种蛋白质，由136个氨基酸组成，分子量15.5 kDa。SAK本身并不是一种酶，它需要与血浆中的纤溶酶原形成复合物而起到溶栓作用，若无纤维蛋白或血栓存在，复合物很快被α_2-抗纤溶酶灭活。重组SAK已经批准用于急性心肌梗死的治疗。大多数病人用药后2周内出现高滴度的中和性抗体，但未见严重的过敏反应。临床上常与肝素、阿司匹林联用。

4. bat-PA　吸血蝙蝠唾液中含有一组共4种纤溶酶原激活剂，其中含477个氨基酸的DSPA α_1与t-PA分子具有85%的同源性。哺乳动物细胞表达的重组bat-PA对血栓中纤维蛋白的选择性强，溶栓速度比rt-PA更快，目前正处于临床研究阶段。

5. R-PA　R-PA（reteplase）为野生型t-PA经缺失突变（缺失整个指形区、生长因子区和K_1三角区等结构域，只保留重链的K_2区和轻链结构域的编码序列）后重组表达的单链非糖化分子，可用大肠杆菌表达，成本明显降低，其有效剂量为rt-PA的1/5。美国食品药品管理局（FDA）已批准上市。

6. 水蛭素　水蛭素（hirudin）具有特异的直接抑制凝血酶的功能。天然水蛭素含65或66个氨基酸，N端含有3对二硫键，C端9个氨基酸中有5个为酸性氨基酸，是水蛭素与凝血酶作用的重要功能结构域。肽链中部有一个Pro-Lys-Pro序列，不易被蛋白酶水解。通过基因工程手段获得了重组水蛭素及其系列衍生物，如去硫酸水蛭素（63位Tyr上无硫酸）、S水蛭素（hirugen，人工合成的相当于天然水蛭素C端53～64位12个氨基酸的片段）、Hirulog（一种由连接序列将两条肽链连接起来的分子）等。国产重组水蛭素已批准进入临床试验。

7. 其他在研溶栓药物　其他正在研究中的新型溶栓药物有：重组t-PA突变体（TNK-t-PA，nPA）、重组scu-PA突变体（rscu-PAK158A，rscu-PAK158G，rscuPA-32K，R179-S184缺失突变体等），重组纤溶酶原激活剂的嵌合体（选择性地将各种纤溶酶原激活剂的不同功能区的编码序列进行组合），导向溶栓药物（将t-PA或scu-PA与抗纤维蛋白、抗血小板或抗凝血酶的mAb共价连接）。研制新型溶栓药物的目的就是降低药物的副作用，增强对纤维蛋白的选择性，延长半衰期，拮抗PAI-1的抑制作用，并提高溶栓药物的导向性。

五、其他活性蛋白类药物

1. 可溶性受体和黏附分子（CD分子）药物　某些细胞受体存在膜结合和可溶性两种

形式，前者跨膜存在于细胞膜上，膜外区是其与相应配体结合的结构域，后者是前者经酶切后的膜外区域，游离存在于血浆中。重组可溶性受体是利用基因工程技术，在工程细胞表达受体的膜外区，保留与配体特异性结合的能力，但丧失信号转导功能，并能与膜表面竞争结合配体，从而能发挥受体阻断剂的作用。

细胞黏附分子（cell adhesion molecules）是指由细胞产生、介导细胞与细胞间或细胞与基质间相互接触和结合的分子。黏附分子以配体-受体相对应的形式发挥作用，从而在细胞的信号转导、生长、分化、活化以及炎症、血栓形成、肿瘤转移、创伤愈合等生理和病理生理过程中发挥重要的作用。在某些情况下，特别是疾病状态下，一些细胞表面的黏附分子可以从细胞膜上脱落下来，成为可溶性的黏附分子。利用基因工程技术，可将黏附分子在体外表达成可溶性分子，发挥细胞黏附的阻断作用，治疗因黏附分子功能异常导致的疾病。

利用重组可溶性受体和黏附分子封闭其配体作用的治疗策略已越来越受到人们的重视。与抗体技术相比，重组可溶性受体和黏附分子的主要优点是其属于人体自身成分，在治疗中不易引起免疫应答反应，因而安全性较好。目前批准上市的有可溶性 TNF 受体 2-Fc 融合蛋白（类风湿性关节炎）；其他正处于临床研究阶段的有：可溶性 TNF 受体 1-Fc 融合蛋白（感染性休克、类风湿性关节炎、多发性硬化症）、可溶性 CD54（感冒及自身免疫病），可溶性 CR1/CD21（肺移植导致的再灌损伤、儿科心脏手术）、可溶性 CD40 配体（晚期肾癌）、可溶性 IL-4 受体（哮喘）、可溶性 LFA-3/IgG1 融合蛋白（银屑病）等。

2. 抗菌肽类药物　杀菌性/通透性增强蛋白（bactericidal/permeability-increasing protein，BPI）是从多形核粒细胞中分离出的一种阳离子蛋白质，分子量为 55 kDa，具有特异性杀伤革兰阴性杆菌和中和细菌内毒素的作用。氨基酸序列分析发现 BPI 的杀菌活性、中和 LPS 和结合肝素的活性区域分别位于第 17～45、65～99 和 142～169 位氨基酸残基，故认为 BPI 由 3 个独立的功能区组成，这些功能区单独或共同作用的活性是 BPI 生物学效应的基础。

人可溶性 BPI、重组 BPI 以及它的 N-端 25 kDa 片段均可抑制内毒素介导的毒性反应。如能显著降低内毒素攻击小鼠的死亡率，抑制内毒素的鲎试剂反应，抑制内毒素刺激中性粒细胞表达 CR 和 CR3，对内毒素诱导的纤维蛋白溶解和凝血系统活化具有明显的保护作用，抑制内毒素介导的内皮细胞损伤和 IL-6、TNF-α 和 IL-1β 等的释放等。

重组 BPI 21 和 BPI 23 在Ⅱ-Ⅲ期临床研究中显示出良好的效果，对创伤后出血性休克、肝部分切除术后、肠源性内毒素移位、感染性休克有较好的治疗作用。美国 Xoma 公司用哺乳动物细胞 CHO 表达的 BPI 有 193 个氨基酸，且 132 位的 Cys 突变为 Ala 以防止二硫键错配和分子间多聚体的产生。该重组产品已上市并被用于治疗脑膜炎菌血症、创伤性失血、严重腹部感染和眼科疾患。

其他如阳离子抗菌小肽，包括人防御素、贻贝抗菌肽、爪蟾抗菌肽、乳链菌肽等均得到了重组表达。

3. 骨形成蛋白 2　骨形成蛋白 2（bone morphogenetic protein 2，BMP-2）属转化生长因子 β（TGF-β）家族成员。人天然 BMP 存在于骨基质中，可以 BMP-1、BMP-2A、BMP-2B 和 BMP-3 四种形式存在。BMP-2 的 cDNA 全长 1 587 bp，编码 386 个氨基酸的蛋白质，包括 N-端的信号肽，分子中部的前肽和位于 C-端的成熟肽。BMP-2 在骨科领域可应用于脊柱融合以及大块骨缺损、难愈合性骨折和股骨头坏死等的治疗；在口腔科的应用范围包括牙周再造、牙髓组织修复、根管疗法和颅面颌骨再造等。重组 BMP-2 一般采用 CHO 细胞表达。2002 年 6 月美国 FDA 批准重组 BMP-2 用于脊柱退行性病变的治疗。

第三节　基因药物

基因药物（gene medicines）包括两方面的内涵，一是指将具有治疗意义的DNA重组入真核表达载体，直接转移至人体细胞，在体内表达出具有治疗作用的多肽和蛋白质，这实际上就是基因治疗（见第二十三章）；二是应用直接作用于细胞内DNA或RNA、抑制基因复制和表达的核酸片段或其人工合成的类似物，以及对RNA具有特异切割作用的核酶等。所以从基因药物的作用机制来讲有两种情况，一是导入完整的外源性功能（治疗）基因，达到基因补充或替代的目的；二是应用与靶DNA或RNA互补的外源性核酸片段或其人工合成产品，干扰基因的复制和表达，又称反义药物，包括反基因药物、反义RNA、小干扰RNA、肽核酸以及对RNA具有特异切割作用的核酶等。

反义技术（antisense technology）是指人工合成或生物合成的DNA或RNA，它们能与DNA或RNA中的靶序列互补，封闭基因的表达，具有特异性强且操作简单的特点，可用于治疗由于基因过度表达所致的肿瘤、遗传性疾病和病毒感染。福米韦森（formivirsen，Vitravene）是1998年美国FDA批准的用于治疗AIDS所致巨细胞病毒视网膜炎的第一个反义药物。国内研制的以端粒酶为靶点的反义核酸抗肿瘤药物正在进行临床前研究。

1. 反基因药物　亦称反基因，实际上为反义DNA。生物体内降解或人工合成的DNA片段进入细胞后与细胞基因组DNA双螺旋大沟中的互补碱基（须包含一段聚嘌呤或聚嘧啶）形成碱基三联体如T-A＊T，C-G＊C，进而形成局部DNA三螺旋。因此，反义DNA又称三链形成寡核苷酸（triplex-forming oligonucleotides，TFO）。由TFO和细胞内DNA形成的三螺旋结构可抑制序列特异的DNA结合蛋白或转录因子的结合，从而干扰DNA的复制和转录。例如，研究发现三螺旋的形成抑制了转录因子与转录起始位点上游调节序列的结合，因而阻断c-*MYC*的转录；三螺旋的形成也可抑制培养细胞IL-2受体α亚基mRNA的合成。也有一些TFO可形成夹钳结构，亦能与特定序列的RNA结合形成DNA-RNA杂交体，从而抑制特定基因的表达。

识别并能与细胞内基因组中DNA的某一靶序列形成三螺旋结构的TFO最小长度为17 bp。为了能形成更稳定的三螺旋结构，在合成TFO时可共价联结嵌入剂。TFO经修饰后可抵抗核酸酶的作用。

2. 反义寡核苷酸　反义寡核苷酸（oligodeoxynucleotides，ODNs）实际上是人工合成的反义单链DNA（cDNA），作用的目标分子主要为mRNA。

针对不同靶序列的ODNs可发挥不同的作用。与双链DNA结合后可调节转录，与mRNA前体或剪接点结合可抑制mRNA前体的剪接；互补的ODN-mRNA杂交体影响剪接后mRNA从胞核转运至细胞质；与细胞质中的mRNA结合可阻断翻译；与某些特定的蛋白质结合还能调节基因表达。此外，反义ODNs与目标mRNA形成杂交体后，易促进RNase H的活性而使杂交体中的RNA被降解。反义ODNs对病毒的反转录过程亦可发挥作用，可通过引物竞争、cDNA延长终止，以及与多聚酶结合达到抑制作用。

反义ODNs具有毒性小、易吸收、与靶分子杂交迅速、诱导RNase H活力较强等特点，但分子不稳定，易被体内的核酸酶降解。另外，ODNs上的负电荷阻碍其通过细胞膜。为此，须对ODNs进行分子修饰，方可提高其效果。常用的修饰方法为甲基或硫原子修饰。甲基磷酸型ODNs（磷酸根上的羟基被甲基化）稳定，易透过细胞膜，但杂交率低，诱导RNase H活力差；硫代磷酸型ODNs（磷酸根上的1个羟基被巯基取代）稳定，呈脂溶性，更易透过细胞膜，诱导RNase H活力强，其缺点是吸收差且具有细胞毒性；

双硫代磷酸型 ODNs（磷酸根上的 2 个羟基均被巯基取代）抗核酸酶活性强，但合成复杂，且具有细胞毒性。除此之外，在实际应用时还可在 ODNs 组方中加入一些化学试剂，以促进其透过细胞膜或增加其对 RNA 的亲和力。

3. 反义 RNA　反义 RNA（antisense RNA）指能与靶分子 mRNA 序列互补的 RNA 片段，通过碱基对之间氢键的作用与靶 mRNA 形成双链复合物，影响基因的转录、翻译和加工过程，从而调控基因的表达。从理论上讲，反义 RNA 应该属于负调控方式，起作用的环节是转录后 mRNA 从胞核到胞浆的转移阶段和翻译起始阶段。

反义 RNA 可直接人工合成，亦可通过合成反义 RNA 的基因（cDNA），构建表达载体，导入合适的宿主细胞后，进行大量的制备。针对靶 RNA 不同区域序列设计的反义 RNA 可发挥不同的抑制活性（表 22-3）。

表 22-3　反义 RNA 的抑制活性

针对靶 RNA 不同区域序列的反义 RNA	作用机制	表现活性
引物 RNA	反义 RNA-引物 RNA 复合物	抑制 DNA 复制
mRNA 5′端编码区	反义 RNA 结合起始密码子 AUG	抑制 RNA 转录，抑制翻译启动
mRNA 外显子与内含子连接部位	反义 RNA 结合 GTAG	干扰前 mRNA 拼接
mRNA 3′端非编码区	反义 RNA 结合 poly（A）	抑制 mRNA 的成熟和向胞浆的转运
mRNA 5′端非编码区	反义 RNA 结合后阻断 Cap 结构的形成	抑制 mRNA 的成熟和向胞浆的转运
mRNA 5′端非编码区	反义 RNA 结合 SD 序列	抑制蛋白质翻译

反义 RNA 易获得，作用的靶位特异性强，不改变基因的结构，不会引起基因整合，故希冀应用于病毒病和肿瘤的治疗。例如针对反式激活蛋白反应顺序（TAR）的反义 RNA 可用于抑制 HIV 基因的转录，针对原癌基因 *c-MYC*、*RAS*、*FOS*、*JUN*、*ABL* 等的反义 RNA 可用于恶性肿瘤的治疗。但上述方案均处于实验阶段。

4. 小干扰 RNA　治疗用小干扰 RNA（siRNA）是指人工合成的短（21～23 nt）双链 RNA，可抑制细胞内特定基因的表达，导致转录后基因沉默（PTGS）。根据 RNAi 的基本原理，可人工合成 siRNA，有目的地干扰特定基因的功能，从而在基因的功能研究和开发以 RNAi 为基础的治疗药物方面提供了新的思路和技术路线。

5. 肽核酸　肽核酸（peptide nucleic acid，PNA）是一种以肽为骨架的 DNA 类似物，它是在反义脱氧寡核苷酸的磷酸-脱氧核糖骨架以肽键取代后得到的一类新型化合物。PNA 以 2-氨基乙基甘氨酸为骨架，4 种碱基为侧链，碱基通过亚甲羰基连于骨架上，保持与天然单链 DNA 分子中相邻碱基间（6 个键）、碱基与主链骨架间（2～3 个键）相近的链数目。

PNA 保留了与 DNA、RNA 杂交的能力，而且由于其不带负电荷，因此与 DNA、RNA 的亲和性更高。它能像正规的寡核苷酸一样区分错配碱基，且能抵抗核酸酶的降解作用。PNA 能与双链 DNA 结合，但不形成 PNA-DNA-DNA 三联体。PNA 通过链置换（strand replacement）与双链 DNA 中的一条链的局部发生结合，形成 D 环。PNA 与靶 DNA 结合后，对转录可发挥正负两方面的调节作用，分别实现对转录的序列专一性的终止或者启动。如 PNA 与处于模板链上较短的靶序列相互结合即可实现对转录的抑制作用；而与非模板链上的靶序列结合后，能引起与之对应的模板链上基因的开放，从而实现对转录的正调控。PNA 亦能与 mRNA 中的互补序列结合，产生序列专一性的翻译抑制作用。

6. 核酶　核酶又称催化 RNA，俗称“基因剪刀”。它是天然的具有催化能力的 RNA 分子，在许多生命过程中起着重要作用，如 RNA 的自我裂解、自我剪接、tRNA 的转录后加工，以及染色体端粒的合成等。

经过基因工程改造的核酶可以位点特异性切割任意给定的 RNA 分子，它可识别 mRNA分子特异结构部位，将 mRNA 链剪断，从而抑制基因的表达。

核酶与反义 RNA 一样，对靶向 mRNA 具有高度的选择性。由于核酶是催化性分子，一个核酶分子可以裂解与之结合的靶向 mRNA，将其释放，再裂解其他靶向 RNA，因此具有催化放大的效应，实际应用时使用量低，这一点明显优越于反义 RNA，因此特别适用于基因组变异快速病毒的治疗。可设计针对多种病毒 RNA 靶向序列的核酶，同时用于裂解病毒 RNA，病毒就较难产生耐受性。

核酶可应用于恶性肿瘤和病毒感染的治疗。例如，针对 *RAS* 的核酶可用于黑色素瘤的治疗，针对 *BCR-ABL* 的核酶可用于髓性白血病的治疗；针对 HIV 反转录酶和蛋白酶 mRNA 的核酶可用于 AIDS 的治疗；针对 *MDR-1* 的核酶可提高化疗药物对肿瘤的敏感性；针对基质溶解酶的核酶可以用于治疗关节炎等。但大多数方案仍处于研究阶段，尚未见实际应用的报道。

第四节　基因工程疫苗

早在 12 世纪我国就发明了疫苗，但直到 1798 年才由 E. Jenner 将无害的牛痘疫苗用于天花的预防，并在全世界推广并获得巨大的成功。200 多年来，科学家研制的各种疫苗在控制传染病方面发挥了重要的作用，在一些国家，脊髓灰质炎、破伤风、白喉等疾病的发病率已经很低，天花已在全世界范围内消灭。同时由于抗菌药物的不断创新，使得细菌的感染和死亡率大幅度下降。但由于抗生素的滥用，使得细菌获得了耐药性，并且产生了一些新的突变菌株，给临床治疗带来了新的困难。为此，研究和开发新的疫苗，特别是基因工程疫苗，对于传染病和肿瘤等的预防和治疗具有重要的现实意义。

传统疫苗主要包括减毒活疫苗、灭活疫苗和用某些抗原成分制成的亚单位疫苗。现代生物技术的发展为疫苗的研制带来了新的机遇，人们可以通过重组 DNA 技术克隆表达保护性抗原基因，利用表达产物或重组体制成疫苗，这就是基因工程疫苗（gene engineered vaccine）。基因工程疫苗主要包括基因工程亚单位疫苗、基因工程载体疫苗、核酸疫苗、基因缺失疫苗和蛋白质工程疫苗；按照使用目的来讲可以分为基因工程病毒疫苗、细菌疫苗、寄生虫疫苗和肿瘤疫苗等。

一、病毒疫苗

1. 基因工程亚单位疫苗（gene engineered subunit vaccine）　即纯化的基因工程表达的蛋白抗原，这种疫苗的免疫原性好，可消除病毒潜在的致癌性，并可大量获得高纯度的疫苗。目前高效价廉的基因工程乙肝疫苗已取代传统的血源乙肝疫苗。这种疫苗具有良好的安全性，但其免疫效果有待进一步提高。目前正在研究的基因工程亚单位疫苗主要包括甲肝、丙肝、戊肝、EBV、出血热和艾滋病等的疫苗。

2. 载体疫苗（vectored vaccine）　将保护性抗原基因重组到减毒或无致病性的微生物（载体）中，当这类微生物感染机体后可表达相应的保护性抗原。这类疫苗多为活疫苗，免疫接种后重组体可在体内大量增殖产生目标抗原，从而使机体连续产生特异性免疫保护反应，并且载体本身还有免疫增强的作用。常用的载体包括痘病毒、腺病毒载体以及卡介苗等。目前以痘病毒为载体的疫苗如甲肝、乙肝、HIV 等重组疫苗已进入少量人体观察阶段。

3. 核酸疫苗（nucleic acid vaccine）　亦称基因疫苗（gene vaccine），是以编码保护性

抗原的基因而制成的疫苗。小鼠直接注射含有抗原基因的质粒或 RNA 均可使机体产生相应的抗体。自 1992 年以来，已经有多种核酸疫苗在进行研究，由于 RNA 容易降解，一般使用 DNA 疫苗，其优点在于易于制备、便于保存、可多次免疫，并容易制成多联多价疫苗。

4. 基因缺失活疫苗（gene deleted live vaccine）通过基因缺失突变使得病毒的毒力降低或消失，而不引起抗原性改变所制备的基因工程病毒活疫苗。这种疫苗突变性状明确、稳定、不易返祖。如目前使用的已去除了与毒力相关基因的兽用伪狂犬病毒疫苗。

5. 蛋白质工程疫苗（protein engineered vaccine）使抗原基因发生突变、抗原构型改变，或将不同基因或结构域融合，使得表达产物的免疫原性增强，副作用降低的一类疫苗。

二、细菌疫苗

传统的细菌疫苗包括：通过加热或化学处理得到的灭活菌苗、传代培养获得的减毒活菌苗、从细菌培养液中提取纯化得到的亚单位或成分疫苗等。这些疫苗存在效价不够高，需要多次免疫等缺点。随着现代分子生物学、现代微生物学与免疫学的发展，目前可以通过构建突变株、构建重组细菌载体、人工合成多肽和制备核酸疫苗等途径来获得新的细菌疫苗。

1. 霍乱菌苗　霍乱弧菌可分为 200 个以上的血清型。近 10 年来研制开发的新霍乱菌苗包括：①　BS/WC 灭活菌苗：由提取的或重组的霍乱毒素 B 亚单位及灭活的菌细胞组成，6 个月内保护效果达 85%；②CVD103-HgR 减毒活菌苗：经典型霍乱菌株 569B 构建而来，并在染色体上插入了筛选标记。经Ⅰ、Ⅱ期临床观察，服用安全，有效保护率达 88%；③伤寒/霍乱双价菌苗：是以沙门菌为载体表达霍乱弧菌 O 抗原的双价菌苗，对霍乱的保护效率不高（25%）；④O139 减毒菌苗：候选株 CVD112 和 Bengal-15 的有效保护率可达 80%。

2. 痢疾菌苗　痢疾是我国常见的多发性传染病，随着对痢疾致病性和抗痢疾免疫防御机制研究的发展，痢疾菌苗从灭活的死菌苗、提取的成分苗、非侵袭性活菌苗发展到通过基因工程制备的侵袭性活菌苗。

20 世纪 80 年代以来我国一直致力于开发双效价痢疾疫苗，牟兆钦等将编码宋内菌 LPS-O 抗原的质粒转移至 T32 菌中，获得了 FSM2117 和 FS-18 痢疾菌株，宋树珍等在此基础上构建了 FS 双价痢疾菌苗，经观察，这两个菌株均安全有效，已获得国家新生物制品证书和试生产文号。目前国内外正在研制第三代痢疾菌苗，即通过失活或缺失的方法构建减毒突变株，以保存菌苗的入侵能力，但降低其无限繁殖能力。

3. 伤寒杆菌疫苗　伤寒杆菌可入侵肠黏膜上皮细胞，经淋巴入血后引起严重的菌血症。死菌苗的应用虽然已有 70 余年的历史，但由于副作用大而推动了活菌苗的研制。目前有 Ty 21a 和伤寒 Vi 多糖菌苗两种活菌苗在使用。Ty 21a 是伤寒杆菌 VigalE 经 NTG 诱导后的减毒株，缺乏 UDP-4-半乳糖异构酶而不能合成细胞壁脂多糖，有 87%的保护效果，1989 年 FDA 批准上市。伤寒 Vi 多糖菌苗，则是通过提取位于菌膜 LPS 外的荚膜多糖抗原 Vi 制备的。

4. 幽门螺杆菌疫苗　幽门螺杆菌是引起慢性胃炎和消化性溃疡的致病菌，也是胃癌的诱因，世界上超过 50%的人曾感染该菌。幽门螺杆菌可与保护胃黏膜上皮细胞的层粘蛋白的 N-乙酰神经氨酰乳糖特异结合，破坏胃黏膜上皮细胞的完整性。用幽门螺杆菌的粗制抗原与霍乱毒素可以有效阻止幽门螺杆菌感染小鼠的胃。目前重组脲酶与大肠杆菌热不

稳定毒素（LT）融合蛋白的疫苗已在美国和瑞士进行临床观察。

5. 结核杆菌疫苗　结核杆菌是需氧胞内致病菌，菌体表面含有多种不同的蛋白、糖类和脂类的大分子化合物。卡介苗（BCG）是当前预防结核病的唯一活菌苗，早在20世纪即得到广泛认可和应用，但艾滋病患者和HIV携带者接种BCG后出现全身感染，促使人们研制新一代结核菌苗。目前结合杆菌的全基因组测序已经完成，这将大大促进核酸菌苗或其他结核菌苗的研制。

6. 肠出血性大肠杆菌菌苗　1982年发现的大肠杆菌O157：H7是肠出血性大肠杆菌的首要病原菌，1996年日本大规模暴发的出血性肠炎之罪魁祸首就是该菌。O157：H7通过特殊的黏附分子黏附于靶器官，产生类志贺毒素（Slt）与肠溶血素（hly），引起出血性腹泻和溶血性尿毒综合征等。目前已进入临床观察的菌苗有O157多糖与绿脓杆菌外毒素蛋白的结合菌苗（O157LPS-rEPA）。

三、寄生虫疫苗

1. 疟疾疫苗　疟疾是世界上流行最广的寄生虫病。由于疟原虫的抗药性增加以及传播疟疾的蚊子对杀虫剂的抗药性，使疟疾在世界范围内持续流行。我国自主研制的重组疟疾疫苗PfCP-2是将恶性疟原虫AMA-1（Ⅲ）和MSP1-19两个抗原蛋白连接而成，并以酵母为载体高效表达新的疫苗抗原。该疫苗经过动物试验，证明有显著抑制恶性疟原虫生长的功效，目前已进入临床研究。英国葛兰素史克公司的科学家开发出一种新型疟疾疫苗，已在2000名2至4岁的儿童群体中进行试验。此外国际上正在进行疟疾核酸疫苗的研制。

2. 血吸虫病疫苗　血吸虫病是一种人畜共患性慢性寄生虫病。10多年前发明的口服新药吡喹酮治愈了很多血吸虫病患者，但在停药后，他们中却有70%的重复感染。所以预防成为控制血吸虫病的首选措施。目前已鉴定出很多对实验血吸虫病有保护性的抗原，这是研制血吸虫病疫苗的基础。国内外正在进行基因工程血吸虫疫苗及核酸疫苗的研究，其中曼氏血吸虫抗原Sm480是一种较好的候选疫苗分子。

四、肿瘤疫苗

肿瘤疫苗（cancer vaccines）是利用肿瘤细胞或肿瘤抗原物质诱导机体的特异性细胞免疫和体液免疫反应，以增强机体的抗癌能力，阻止肿瘤的生长、扩散和复发，因此又称为肿瘤特异性主动免疫治疗。随着分子生物学理论的研究和生物工程技术的进步，新的肿瘤疫苗不断出现，有些已获准进入临床试验阶段。

1. 肿瘤细胞及其衍生物疫苗　即原始的肿瘤疫苗，它是以灭活的自身肿瘤细胞或其粗提物作为抗原，加佐剂后进行免疫治疗。

2. 肿瘤亚单位疫苗　应用分子生物学方法制备出肿瘤相关抗原（TAA）、肿瘤特异抗原（TSA）以及肿瘤抗原肽疫苗来进行免疫治疗。

3. 肿瘤基因疫苗　这是通过基因重组技术，将目的基因导入受体细胞所制备的疫苗，也是目前发展最快、备受人们重视的一个研究领域。这种疫苗具有以下优点：①可以提高机体抗瘤能力：如将某些免疫增强基因导入免疫细胞，使这些外源基因在免疫细胞中表达，再回输到病人体内，这样就可加强机体免疫系统的功能，特别是使输入的免疫活性细胞在肿瘤局部区域分泌出高浓度的细胞因子，以促进对肿瘤细胞的杀伤。如携带IL-15编码基因的瘤苗可提高T细胞对抗原的反应，编码IL-18的瘤苗能介导CD_4^+ T细胞和NK细

胞的抗瘤效应。②增强肿瘤细胞的免疫原性：将可增强肿瘤细胞抗原性的有关基因（如B7基因）导入肿瘤细胞，使免疫细胞易于识别，不能逃避免疫监视并促进细胞毒性T淋巴细胞的抗瘤活性。③表达产物的直接杀瘤活性：如将*TNF*基因导入肿瘤浸润淋巴细胞或肿瘤细胞，使之能在体内持续分泌TNF，以此杀伤肿瘤细胞。

4. 核酸疫苗　由能引起保护性免疫反应的抗原基因片段及其载体建构而成。包括DNA疫苗和RNA疫苗，目前研究较多的是DNA疫苗。核酸肿瘤疫苗能同时激发机体细胞和体液免疫反应，并诱导许多亚单位疫苗不能诱导的CTL效应。

5. 抗独特型抗体疫苗　抗独特型抗体具有模拟抗原及免疫调节的双重作用，能克服机体免疫抑制，打破免疫耐受，故能代替肿瘤抗原诱导特异性主动免疫反应。目前，针对黑色素瘤、结直肠癌、卵巢癌、B细胞淋巴瘤、皮肤T细胞淋巴瘤和乳腺癌的抗独特型抗体的研究取得了一定的进展。

第五节　基因工程药物与疫苗的发展前景

21世纪是生命科学的世纪，也是以基因工程为关键技术的生物制药飞速发展的世纪，为此，世界各国都制订了详尽的发展规划，力争在国际重组药物市场中占有一席之地。我国重组药物的研究与开发在国家“863”计划、科技攻关（支撑）计划、“火炬计划”和“攀登计划”等的支持下于20世纪90年代起迅速发展，目前已有30余种重组药物上市，随着功能基因组学研究的不断深入，将有数千个功能基因可用作新药的开发。

目前，重组药物的研制主要针对一些严重威胁人类健康的重大疾病，如癌症、艾滋病、自身免疫性疾病、心脑血管疾病、神经障碍性疾病、呼吸系统疾病、糖尿病和遗传疾病等。我国加入WTO后，重组药物首先要面对的是知识产权的问题，而此前我国只有屈指可数的重组药物具有自主知识产权。因此，应重视和加强以下领域的研究和应用。

（1）从天然产物中筛选可供基因重组的药物：在天然活性物质研究方面，我国有一定的基础，从动物、植物和微生物中发现和鉴定了一批具有药用价值的蛋白质成分。当前，各国正在加紧海洋药物的开发和利用，我国有丰富的海洋资源，从深海或极端环境中提取活性物质，并对其进行人工改造和重组表达是研发创新药物的一条捷径。

（2）对现有药物进行蛋白质工程改造，获得疗效更好、靶向性更明确的新药：充分利用现有资源，对有治疗意义的活性蛋白或多肽进行分子修饰，一方面可避免药物开发的重复性和知识产权，另一方面可提高现有重组药物的疗效和靶向性。如TNFα由于严重的毒副作用如低血压和肝毒性，因而不能作为全身用药，通过蛋白质工程技术可对其结构的某些部位进行氨基酸置换，从而得到疗效提高、毒性降低的衍生物。

（3）加强基础和应用基础研究，通过功能基因组学的研究挖掘新基因、新蛋白：通过生物信息学平台，充分利用我国人口众多、疾病资源丰富的优势，从浩如烟海的数据库中挖掘开发新基因，新蛋白作为重组药物的候选分子。目前国内的众多大学和生物制药公司都在这方面投入了大量的人力、物力和财力，相信不久的将来我国将会有一批具有自主知识产权的重组药物问世。

（4）应用新的技术与方法，发现新的药物靶标，为新药的源头创新提供理论依据：应将生物信息学、功能基因组学、药物基因组学、蛋白质组学、RNA干扰技术、噬菌体展示文库、组合化学、高通量筛选、计算机辅助设计等新技术、新方法运用于新的药物靶标的寻找和筛选中去，为设计新药提供理论依据和奠定基础。这将是研究和开发重组药物的重要途径。

（5）反义技术和反义药物的应用前景广阔：利用反义技术，通过研究特定基因的表达

与肿瘤发生发展的关系，可揭示肿瘤的病因和致病机制。而通过对特定癌基因的抑制，可以有效地阻止肿瘤的发展，为肿瘤的防治提供新的手段。应用反义药物抑制病毒基因的复制与表达，可探索和阐明病毒的致病机制，并有望发展成新的、有效的病毒感染治疗手段。同样，只要获得合适的靶位，反义技术及其药物在心血管、自身免疫性疾病等的治疗方面也将发挥重要的作用。目前，反义药物已成为国外制药业研究的一个重点和热点，国内应及早跟进和发展。

（焦炳华）

第二十三章　基因治疗

关键知识点

- 基因治疗是以改变遗传物质为基础的生物医学治疗，即通过一定方式将正常基因或有治疗作用的DNA片段导入人体靶细胞以矫正或置换致病基因的治疗方法。目前的基因治疗方案主要包括：将正常基因导入患者细胞内以弥补基因突变或缺失造成的功能缺陷，或抑制有害基因的表达等。
- 将治疗基因导入体内的方法主要有非病毒载体法和病毒载体法。非病毒载体可用直接注射、电穿孔和脂质体转染导入。非病毒基因导入法安全性好但效率低。病毒载体法将野生型病毒改造成为无致病的基因转移载体。目前常用的有反转录病毒、腺病毒、腺相关病毒、单纯疱疹病毒等。它们各具有不同的优缺点，可依据不同的要求加以选择。
- 基因治疗的常用靶细胞有造血干细胞、体细胞和肿瘤细胞等。
- 病因明确的单基因遗传病基因治疗方案容易确定，如血友病。多基因相关疾病如高血压、动脉粥样硬化、心肌肥厚等，基因治疗效果尚不理想。恶性肿瘤的基因治疗方案已有多种尝试，包括：将抑癌基因导入肿瘤细胞中表达，抑制癌细胞生长；利用反义RNA、核酶和小干扰RNA抑制癌基因过表达；将细胞因子基因导入受体细胞，使其持续产生细胞因子而激活和增强机体抗肿瘤免疫反应等。
- 虽然基因治疗取得了重大进展，但仍有诸多问题待解决。基因治疗必须严格遵守国家有关法律规定，经过严格的安全评价，获得批准后方可进入临床试验。

关键词： 基因治疗；基因转移；靶细胞；基因干预

复杂的遗传背景、环境因素和社会因素使得人类遗传性疾病、恶性肿瘤、心脑血管性疾病、糖尿病等疾病的发病率和死亡率日益上升，治愈这类疾病对于医学界是巨大的挑战。基因结构及基因表达调控的异常是上述疾病发生发展的重要原因，通过DNA重组技术纠正这些异常无疑将为这些疾病的治疗带来革命性的变化。基因治疗是这些策略中最直接、最富于挑战性，也可能是最有前景的尝试。第一例真正意义上的基因治疗是用腺苷脱氨酶（ADA）基因来治疗*ADA*基因缺陷引起的严重复合型免疫缺陷症（SCID）的患者，正常*ADA*基因借助反转录病毒载体被导入患儿的T淋巴细胞，再回输给患儿，5年后患儿体内仍有10%的造血细胞呈ADA基因阳性，这一治疗病例的成功使得基因治疗迅速开展起来。

基因治疗（gene therapy）是指改变以人遗传物质为基础的生物医学治疗，即通过一定方式将人正常基因或有治疗作用的DNA片段导入人体靶细胞以矫正或置换致病基因的治疗方法。它针对的是疾病的根源-异常的基因本身。在此过程中，目的基因被导入靶细胞内，它们或与宿主细胞染色体整合成为宿主细胞遗传物质的一部分，或不与宿主细胞的染色体整合而独立于染色体以外，但都可以在宿主细胞内表达基因产物蛋白质，而达到治疗作用。目前基因治疗的概念也有了较大扩展，凡是采用分子生物学技术和原理，在核酸水平上展开的对疾病治疗都可纳入基因治疗范围。

基因治疗的范围也从过去的单基因遗传病扩展到恶性肿瘤、心脑血管疾病、神经系统疾病、代谢性疾病等。

第一节 基因治疗的策略和基本程序

基因治疗的策略和方法在过去的20多年中不断改进和发展，已有相当一部分进入了临床试验阶段。本节讨论基因治疗的目的、可选择的基因、可用于导入基因的载体、导入方式、可导入基因的细胞种类等基本策略和程序。

一、基因治疗可修复错误基因或干预异常表达的基因

基因治疗的基本方式主要有以下几类：

1. 基因置换 用正常的基因对基因组中错误（突变）基因在原位进行基因置换（gene replacement）或基因矫正（gene correction）是最理想的基因治疗策略。对于由单个基因突变引起的遗传病，采取这种方法既不破坏整个基因组的结构，又可达到治疗疾病的目的。但是这种属于基因定点同源重组的技术尚处于基础研究阶段，目前尚未能从理论和技术上得到突破。

2. 基因增补 不删除突变的致病基因，而在基因组的某一位点额外插入正常基因，在体内表达出功能正常的蛋白质，达到治疗疾病的目的是目前主要使用的基因治疗策略。这种对基因进行异位替代的方法称为基因添加（gene augmentation）或称基因增补。基因增补不仅可以用于替代突变基因，也可以在原有基因表达水平不足以满足机体需要的情况下，异位过表达来增强体内某些功能。例如将编码干扰素α或β和IL-2等分子的基因导入恶性肿瘤患者体内，可以激活体内免疫细胞的活力，作为抗肿瘤治疗中的辅助治疗而达到治疗肿瘤的目的。

3. 抑制有害基因表达或过度表达的基因 有些疾病是由于某一或某些基因的过度表达引起的，抑制这些基因的表达就可以治疗这类疾病。因此向患者体内导入有抑制基因表达作用的核酸，如反义RNA、核酶、小干扰RNA、肽核酸等，通过这类核酸类药物抑制或降解相应的mRNA，从而达到治疗疾病的目的。

二、利用载体将目的基因导入到受体细胞表达是基因治疗的基本程序

基因治疗的基本过程包括4方面：①选择治疗基因；②利用载体把目的基因导入到受体细胞表达；③选择基因治疗的靶细胞；④治疗基因表达的检测。

（一）治疗基因的选择

基因治疗是将正常基因代替致病基因，以在细胞内产生有正常功能的蛋白质为治疗的目的。因此，细胞内的基因在理论上均可作为基因治疗的选择目标。许多分泌性蛋白质如生长因子、多肽类激素、细胞因子、可溶性受体，以及非分泌性蛋白质如受体、酶、转录因子、原癌基因产物及抑癌因子等的正常基因都可作为治疗基因。简言之，只要清楚引起某种疾病的突变基因是什么，就可用其对应的正常基因做治疗基因。

（二）利用载体将治疗基因导入细胞或体内表达

大分子DNA不能主动进入细胞，即使进入也将被细胞内的核酸酶水解。因此选定治疗基因后，怎样把治疗基因安全有效地导入细胞内表达是基因治疗的关键。目前已发展的基因转移（gene transfer）系统有两大类：病毒载体系统和非病毒载体系统。

在基因治疗实施中把治疗基因导入病人体内的方式有两种：一种是直接的体内疗法

(*in vivo*)，即将携带有治疗基因的载体直接导入体内有关的组织器官，使其进入相应的细胞并进行表达。另一种是间接体内疗法（*ex vivo*），即先把待接受基因的靶细胞从体内取出，在体外培养，将携带有治疗基因的载体导入培养细胞内，经过筛选把接受了治疗基因的细胞挑选出来，繁殖扩大后再回输体内有关组织中，使治疗基因在体内表达相应产物，以达到治疗的目的。

（三）基因治疗的靶细胞

基因治疗所采用的靶细胞通常是体细胞，包括病变组织细胞或正常的免疫功能细胞。生殖细胞不能作为基因治疗的靶细胞，因为人类生殖生物学极其复杂，目前对许多问题还没有研究清楚，因此基因治疗的原则是仅限于患病的个体，而不能涉及人类的下一代，为此国际上严格限制用人生殖细胞进行基因治疗实验。人类的体细胞有 200 多种，目前还不能对大多数体细胞进行体外培养，因此能用于基因治疗的体细胞十分有限。目前能成功用于基因治疗的靶细胞主要有：

1. 造血干细胞　造血干细胞（hematopoietic stem cell，HSC）是骨髓中具有高度自我更新能力的细胞，能进一步分化为其他血细胞，并能保持基因组 DNA 的稳定。HSC 已成为基因治疗最有前途的靶细胞之一。由于造血干细胞在骨髓中含量很低，难以获得足够的量用于基因治疗。人脐带血细胞是造血干细胞的丰富来源，它在体外增殖能力强，移植后抗宿主反应发生率低，是替代骨髓造血干细胞的理想靶细胞。目前已有脐带血基因治疗的成功病例。

2. 皮肤成纤维细胞　皮肤成纤维细胞具有易采集、可在体外扩增培养、易于移植等优点，是基因治疗有发展前途的靶细胞。反转录病毒载体能高效感染原代培养的成纤维细胞，将它再移植回受体动物时，治疗基因可以稳定表达一段时间，并通过血液循环将表达的蛋白质送到其他组织。

3. 肌细胞　将裸露的质粒 DNA 注射入肌组织，重组在质粒上的基因可表达几个月甚至 1 年之久，这对于 Duchenne 肌营养不良症（DMD）是一个理想的基因治疗方案。由于肌细胞有特殊的 T 管系统与细胞外可直接相通，利于注射的质粒 DNA 内吞进入细胞内。而且肌细胞内的溶酶体和 DNA 酶含量很低，环状质粒在胞浆中存在而不整合入基因组 DNA，能在肌细胞内较长时间保留，因此骨骼肌细胞是基因治疗的很好靶细胞。

可用于基因治疗的靶细胞还有血管内皮细胞、神经细胞、上皮细胞以及肿瘤细胞等。

三、病毒载体和非病毒载体均可用于导入目的基因

目前用于基因治疗的载体有两种：病毒载体和非病毒载体，所以基因转移的方法也分为病毒法和非病毒法，即生物学法和非生物学法。非生物学法主要可分为物理和化学法。物理和化学方法操作简单、安全，但效率不高。病毒学方法的特点是基因转移效率高，但安全问题需要重视。

（一）基因转移的非生物学方法

基因治疗中所用的治疗基因，采用人工哺乳动物细胞表达载体携带，相对于使用病毒载体，称之为非生物学方法。该方法较病毒载体安全，但不能主动有效地进入细胞，转移效率低。

1. 基因转移的物理方法

（1）直接注射法：亦称为裸 DNA 注射法，是将携带有目的基因的非病毒真核表达载体（多为质粒）溶液直接注射入肌组织，目的基因表达时间可长达至一年以上。该法优点

是对机体无毒无害，操作简便。但此法仅限于在肌组织中表达，并需要注射大量的DNA，转入效率低也是难以克服的困难。

（2）基因枪法：又被称为生物弹道技术（biolistic technology）或微粒轰击技术（particle bombardment technology）。其原理就是采用一种微粒加速装置，使裹着外源基因的微米级的金或钨（比重大，且化学性质很稳定）颗粒获得足够的能量后，直接进入靶细胞或组织。基因枪应用于基因治疗具有操作简便、DNA用量少（仅为肌注射用量的1/250～1/2 500）、效率高、无痛苦、适宜在体操作，尤其适于将DNA疫苗导入表皮细胞，获得理想的免疫反应，是基因免疫以及基因治疗领域很具潜力转移方法。但不易于内脏器官的在体操作。

（3）电穿孔法：电穿孔是在直流脉冲电场作用下使细胞膜出现105～115 nm的微孔，这种通道能维持几毫秒到几秒，在此期间质粒DNA可通过通道进入细胞，然后自行恢复。电场取消后不会对细胞造成任何影响。活体电穿孔法（*in vivo* electroporation）可将外源基因选择性地导入靶组织或器官，效率较高，并且可在多种组织器官上应用；对导入的外源基因片段也没有严格限制（100～2 000 kbp）。但也存在外源基因表达持续时间短的缺点。

2. 基因转移的化学方法 包括磷酸钙共沉淀法、脂质体转染法、DEAE-葡聚糖法等转移方法。目前运用较多的是脂质体介导的基因转移。

脂质体（liposome）是由磷脂和相似的两性脂所形成的微囊泡。脂质体的脂双层在结构上与细胞膜相似，可形成具有双层膜的封闭式粒子，携带极性大分子DNA或RNA穿透细胞膜，进入细胞。根据脂质体包裹DNA的方式不同，有阳离子脂质体、阴离子脂质体等。与病毒载体介导的基因转移相比，脂质体具有以下优点：①与细胞膜融合将目的基因导入细胞后，即被降解，对细胞无毒副作用，可反复给药；②不激活癌基因和产生免疫反应；③与基因的复合包被过程容易；④DNA或RNA均能得到有效保护，不至于在胞内被灭活或被核酸酶降解；⑤操作简单快速、重复性好。该方法被认为是最有前途的基因导入方法之一，尤其是适用于恶性肿瘤基因治疗，现在已被美国癌症协会批准为临床基因治疗的第一方案。但也存在一些不足，如体内基因转染效率低，表达时间短，易被血液中的网状内皮细胞吞噬等。

3. 受体介导的基因转移 受体介导的基因转移是利用细胞表面受体能特异性识别相应配体并将其内吞的机制，将外源基因与配体结合后转移至特定类型的细胞。无论是遗传性疾病还是恶性肿瘤的基因治疗，靶向性是非常重要的，特别是应用到体内时，既要考虑对靶细胞的治疗，又要注意对正常细胞的保护。受体介导的基因转移系统可满足上述靶向性的要求，且还具有如下优点：①DNA不整合到宿主细胞染色体，不需要导入细胞处于分裂期；②对转移的外源基因大小无限制；③导入系统本身无潜在的感染性，具安全性，低免疫原性。但缺点是：①瞬时表达，如体内应用，需重复注射；②表达水平低，因为被细胞内吞的复合体可能在内吞小泡与溶酶体融合后被降解，影响外源基因的表达。某些病毒的外膜蛋白质，如复制缺陷型腺病毒、流感病毒血凝素的膜融合区，可辅助外源基因逃逸溶酶体的降解，从而提高外源基因的表达效率。把这些分子与配体连接，受体介导的基因转移在基因治疗中有较好的优势和发展前景。

（二）基因转移的生物学方法

病毒具有靶细胞定向感染性和寄生性两大特征。DNA病毒和RNA病毒都可以作为基因治疗的载体。野生型病毒基因组的编码区主要为衣壳蛋白、酶和调控蛋白编码，而非编码区中则含有病毒进行复制和包装等功能所必需的顺式作用元件。野生型病毒必须经过改造，以确保其在人体内的安全后才能作为基因治疗的载体。改造的方式是剔除病毒复制必

需的基因和致病基因，形成的空间由插入治疗基因填充。原有必需基因的功能改由辅助病毒或包装细胞提供。野生型病毒插入外源基因后的大小不能超过原基因组大小的105%～110%。经过改造的病毒，保留了感染性，去除了致病性，成为介导治疗基因进入细胞的良好载体。

目前用作基因转移载体的病毒有反转录病毒（retrovirus）、腺病毒（adenovirus）、腺相关病毒（adeno-associated virus，AAV）、单纯疱疹病毒（herpes simplex virus，HSV），以及新近发展起来的埃巴二氏（Epstein-Barr，EB）病毒和痘苗病毒（Vaccinia virus）等。不同类型的病毒载体在应用中具有不同的优势和缺点，可依据基因转移和表达的不同要求加以选择。

1. 反转录病毒载体　反转录病毒是正链RNA病毒，其基因组中有编码反转录酶和整合酶（integrase）的基因。在这些酶的作用下病毒基因组RNA被反转录成双链DNA，然后随机整合在宿主细胞的染色体DNA上，长期存在于宿主细胞基因组中，这是反转录病毒作为载体区别于其他病毒载体的最主要优势。科学家利用这一特性，将反转录病毒复制所需要的基因除去，代之以治疗基因，构建成重组的反转录病毒载体。在目前的基因治疗中，70%以上应用的是反转录病毒载体。

（1）反转录病毒的基因组结构：反转录病毒的基因组是两条相同的单链正义RNA，长度在8～11 kbp之间，其5′端有$m^7G^5ppp^5GM$“帽”结构，3′端有（polyA）尾；紧靠5′端和3′端内侧的是一段长末端重复序列（long termination repeat sequence，LTRs），其中含有启动子、增强子及病毒转录所需要的起始和终止信号；其内部序列主要为编码核心蛋白和内膜蛋白的*gag*基因，编码反转录酶、整合酶和其他酶的*pol*基因及编码包膜糖蛋白和跨膜蛋白的*env*基因（图23-1）。

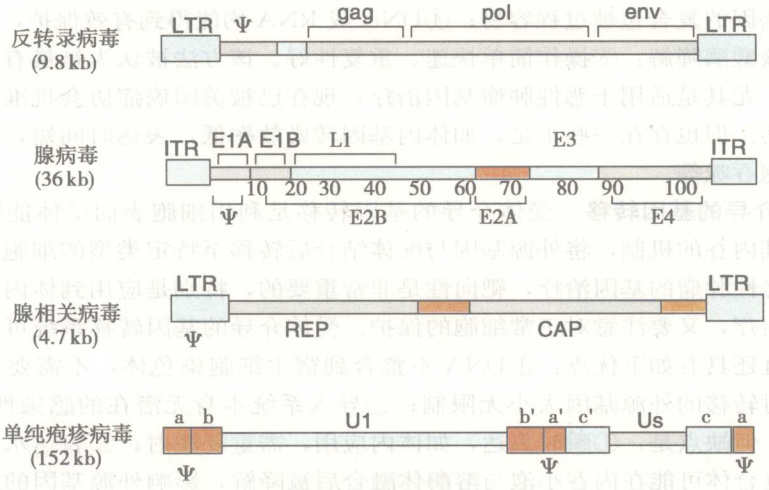


图23-1　常见野生型病毒载体的基因结构

（2）反转录病毒载体的特点：目前基因治疗中采用的反转录病毒载体绝大多数来源于莫罗尼鼠白血病病毒（MoMuLV）。这种反转录病毒能感染鼠、人和其他动物细胞。该病毒通过其外膜蛋白与靶细胞表面的磷酸转运载体结合而进入细胞。在构建反转录病毒载体（图23-2A）时，由于大部分病毒的必需基因和包装信号被治疗基因取代，这样产生的重组前病毒不能表达病毒结构蛋白，因此这种病毒载体已无能力复制和包装成成熟的病毒颗粒。重组的已插入治疗基因的反转录病毒载体必须先在体外一个临时性细胞中复制并包装，获得足量的重组病毒后再用于基因治疗。这种临时性细胞就是包装细胞（packaging

cell)，它是经过特殊改造和修饰的细胞（图 23-2B），已经转染和整合了辅助病毒（helper virus）基因组，可以表达 *gag*、*pol* 和 *env* 等病毒基因编码的蛋白，为反转录病毒载体包装成重组病毒提供全部的病毒蛋白，但它不能转录产生编码完整病毒的基因组 RNA，而且也不产生可能被包装到重组子代病毒中的 RNA，因此，包装细胞本身也不产生任何形式的病毒颗粒。

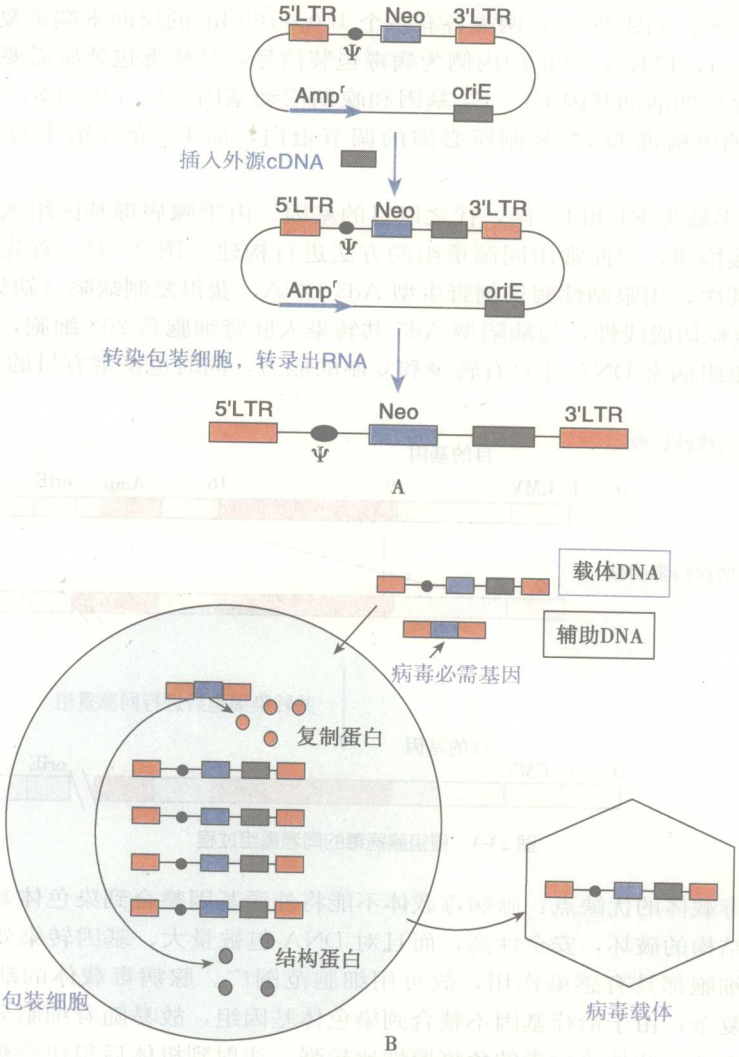


图 23-2　反转录病毒载体的构建和包装

用插入了治疗基因的反转录病毒载体转染包装细胞系，在包装细胞内包装成具有一次性感染能力的病毒颗粒，这种病毒又称为假病毒（pseudo virus）。假病毒即可用于基因治疗。假病毒颗粒中所包含的 RNA 与反转录病毒基因组比较，目的基因和标记基因取代了病毒的结构基因。假病毒感染靶细胞的过程与反转录病毒一样，感染效率非常高，但只能感染处于增殖期的细胞。

（3）反转录病毒载体的优缺点：反转录病毒载体有基因转移效率高、细胞宿主范围较广泛、DNA 整合效率高等优点。缺点主要是在两个方面存在安全性问题，一是病人体内万一有反转录病毒感染，又在体内注射了大剂量假病毒后，就会重组产生有感染性病毒的

可能；二是增加了肿瘤发生机会。后者的原因是由于反转录病毒在靶细胞基因组上的随机整合所致，这种整合可能激活原癌基因或者破坏抑癌基因的正常表达。

2. 腺病毒载体　腺病毒属 DNA 病毒，可引起人上呼吸道和眼部上皮细胞的感染。人的腺病毒共包含 50 多个血清型，并根据其凝血特性分为 A～F 6 个亚类，其中 C 亚类的 2 型和 5 型腺病毒（Ad 2 和 Ad 5）在人体内基本上不致病，适合作为基因治疗用载体。

（1）腺病毒的基因组结构及其作为基因治疗载体的改造：人腺病毒基因组为 36 kbp 双链线状 DNA 分子（图 23-1），两端各有一个 100～160 bp 的反向末端重复序列（inverted terminal repeat，ITR），ITR 的内侧为病毒包装信号，是病毒包装所需要的顺式作用元件。编码区包含早期活动基因 E1～E4 基因和晚期活动基因 L1、L2、L3、L4、L5。其中 E1、E2 和 E4 编码病毒 DNA 复制所必需的调节蛋白；而 E3 介导宿主对腺病毒感染的反应。

Ad 载体大多缺失 E1 和 E3 区，代之以目的基因。由于腺病毒基因组太大、酶切位点复杂，不易直接构建，因此常用同源重组的方法进行构建（图 23-3）。首先构建含有目的基因的质粒；其次，用限制性酶切割野生型 Ad5 DNA，获得复制缺陷（缺失 E1 区）型的 Ad5；然后将质粒切成线性，与缺陷型 Ad5 共转染人胚肾细胞系 293 细胞，只有发生同源重组后形成的重组病毒 DNA 才具有转录和复制的能力，同时也携带有目的基因。

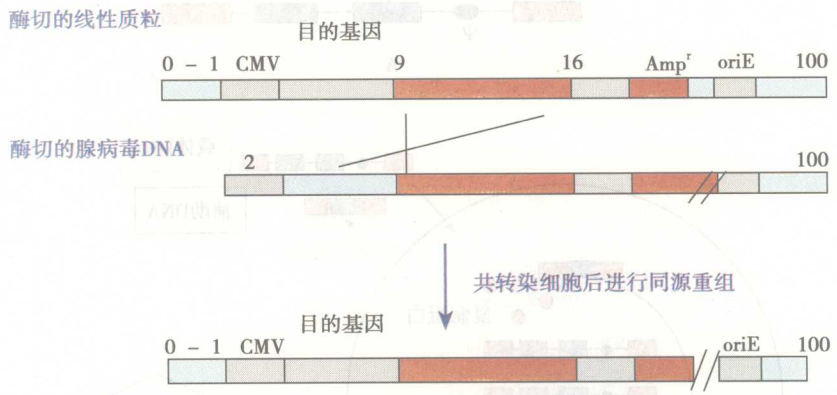


图 23-3　重组腺病毒的同源重组过程

（2）腺病毒载体的优缺点：腺病毒载体不能将外源基因整合到染色体基因组，不会引起病人染色体结构的破坏，安全性高；而且对 DNA 包被量大、基因转染效率高，此外对静止或慢分裂细胞都具有感染作用，故可用细胞范围广。腺病毒载体的缺点是基因组较大，构建载体复杂；由于治疗基因不整合到染色体基因组，故易随着细胞分裂或死亡而丢失，不能长期表达。此外该病毒的免疫原性比较强，注射到机体后很快会被机体的免疫系统排斥。

3. 腺相关病毒载体　腺相关病毒（AAV）是一种细小病毒属的单链 DNA 病毒，病毒基因组长 4 600 bp，包含 2 个基因，*rep* 和 *cap*，分别编码病毒复制和装配必需的蛋白，基因组两端为反向末端重复序列（ITR），是 AAV 病毒复制、整合和包装所需要的顺式作用元件（图 23-1）。用外源基因及其调控序列取代 AAV 的结构基因，保留其两端 145 bp 的反向末端重复序列（ITR），可构建成治疗用重组腺相关病毒载体，在腺病毒的帮助下，包装成为成熟的病毒颗粒。AAV 载体的优点是可以定点整合到 19q13.3，不会引起病人基因组的破坏，不会引起肿瘤；无致病性；载体构建简单，基因表达稳定。缺点是载体本身小，不能转移大的基因。

4. 疱疹病毒载体 Ⅰ型单纯疱疹病毒是一种线状双链 DNA 病毒，基因组长 152 kbp（图 23-1）。HSV-1 是引起人类反复发作性唇疱疹的致病原，它还可以感染并裂解多种类型细胞，并可在某些细胞（如感觉神经元）中潜伏感染，潜伏期可持续终生。因此用它构建载体关键是要考虑安全性。

HSV 载体在遗传病和癌症的基因治疗中应用较多，而且由于 HSV 具有天然的向神经性，因此它在神经系统疾病的基因治疗中具有独特的作用。在帕金森病动物模型中，将酪氨酸羟化酶基因插入到 HSV 载体得到的重组病毒感染纹状体神经元，发现目的基因可以在神经细胞中表达较长时间，但载体的毒副作用较大。将 HSV 的病毒复制相关基因和神经毒性基因突变，得到弱化病毒，这种病毒在正常神经细胞中不复制，而在活跃分裂的肿瘤细胞中则大量复制，并导致细胞裂解。目前，应用 HSV 载体进行的神经胶质母细胞瘤的基因治疗已经进入临床试验。

表 23-1 列举了目前应用和研究较为广泛的几种基因治疗用病毒载体的主要优点和缺点。尽管病毒载体介导的基因转移在临床上已经取得了一定的疗效，但仍存在一些亟待克服的缺陷，包括缺乏有效表达调控、感染的靶向性以及对机体的毒性等。

表 23-1 几类主要病毒载体的特点及其局限性

病毒名称	优　点	不利因素及治疗的副作用
反转录病毒	病毒滴度相对较高； 宿主细胞类型广泛； 整合入宿主细胞基因组，表达长期稳定； 对感染细胞毒性较小； 外源基因插入总容量约 10 kbp	只能感染分裂期细胞； 随机整合可能导致靶细胞基因突变；可通过同源重组产生有复制能力的病毒； 病毒颗粒可被补体途径迅速降解； 可能与人体内源病毒发生重组
慢病毒（HIV 等）	可以感染非分裂期细胞； 宿主细胞类型广泛； 所获病毒滴度相对较高； 整合入宿主细胞基因组，表达长期稳定； 外源基因插入总容量约 10 kbp	可能导致血清转换为 HIV-1 阳性； 病毒调节蛋白可能导致细胞功能失调； 随机整合可能导致靶细胞基因突变； 通过同源重组产生有复制能力的病毒； 可能与人体内源病毒发生重组
腺病毒（Ad）	所获病毒滴度非常高（10^{12} pfu/ml）； 外源基因可得到瞬时的高表达； 可感染包括非分裂期细胞在内多种细胞； 外源基因插入容量大（7～8 kbp 或更大）	容易诱发机体炎症或毒性反应； 免疫原性强，能被抗体所中和； 外源基因不能长期稳定表达； 基因组结构过于复杂
腺相关病毒(AAV)	所获病毒滴度高（10^{10} pfu/ml）； 可感染非分裂期细胞在内的多种细胞； 整合入细胞基因组，表达长期稳定； 定点整合入人细胞 19 号染色体； 不具致病性，毒性小	外源基因容量相对较小（约 2 kbp）； 需在腺病毒或疱疹病毒的协助下包装；重组 AAV 往往失去定点整合的特性；随机整合存在产生插入突变的可能性
单纯疱疹病毒(HSV)	所获病毒滴度在 10^4～10^8 pfu/ml 之间； 外源基因插入长度可达 30 kbp； 可感染神经细胞并使外源基因长期表达； 感染癌细胞并导致细胞坏死	外源基因不能长期稳定表达 容易引起免疫反应、炎症和系统毒性； 基因组结构过于复杂

除上述外，还有一些病毒载体被尝试用于基因治疗。如痘病毒载体，其特点是外源基因的容量大，约为 25 kbp。由肿瘤特异性抗原、细胞因子或病毒致癌蛋白的基因和痘苗病毒参与构建的重组疫苗，在临床应用中能够有效地激发机体的抗肿瘤免疫反应。但是，由于痘苗病毒的结构和生物学特性较为复杂，因此其临床应用的安全性问题尚有待进一步探讨。另外还有辛培斯病毒、人巨细胞病毒、流感病毒等也正在开发作为基因治疗的载体。

近年来，又出现了嵌合病毒载体，即将两种或两种以上的病毒载体加以组合，充分利用这些病毒的优点，可以提高基因转移和表达效率。如腺病毒/反转录病毒嵌合载体等。

四、基因干预可以抑制特定基因的表达

基因干预是指采用特定的方式，抑制某个基因的表达或破坏某个基因的产物，达到治疗疾病的目的。抑制的靶基因往往是过度表达的癌基因或者是病毒复制周期中的关键基因。基因干预作为基因治疗的主要策略之一，越来越受到人们关注。常见的基因干预方法有反义 RNA 技术、核酶技术、RNA 干扰等（见第二十二章）。

第二节　基因治疗的临床应用现状

一、用正常基因替代错误基因治疗遗传病

遗传病的基本特征是由遗传基因改变所引起的。如果只受一对基因影响而发生的疾病叫单基因遗传病，例如镰刀状红细胞性贫血、α-地中海贫血、血友病 A、B，囊性纤维变性等。由多个基因相互作用，并受环境因素影响的人类遗传疾病叫做多基因病，例如：先天性心脏病、唇裂腭裂、原发性高血压、冠心病等。由于这类遗传病由多基因致病而比较复杂，基因治疗的效果还有待于基础研究的突破。

（一）单基因遗传病的基因治疗

单基因缺陷引起的遗传病，由于其致病基因比较清楚，所以基因治疗方案也相对容易确定。基本方案是通过一定的方法把正常的基因导入到病人体内，表达出正常的功能蛋白。例如将人Ⅸ因子基因与反转录病毒重组后转移到 B 型血友病患者自体的皮肤成纤维细胞中，使患者血中Ⅸ因子浓度升高，出血症状及出血次数都明显减少。

（二）代谢性遗传病

由于某种酶蛋白的缺陷，引起由该酶参与催化的物质代谢途径紊乱，全世界发现的这类病有上千种，累计有 200 多种酶的缺陷，其遗传方式多为常染色体隐性遗传，例如苯丙酮尿症、半乳糖血症、次黄嘌呤鸟嘌呤磷酸核糖转移酶缺陷引起的 Lesch-Nyhan 综合征、腺苷脱氨酶（ADA）缺陷引起的严重联合免疫缺陷综合征等，现举 ADA 基因治疗的例子加以说明。

细胞内 ADA 催化腺苷或脱氧腺苷转变为肌苷或脱氧肌苷。当 ADA 缺乏时，细胞内腺苷和脱氧腺苷堆积，进而在激酶作用下磷酸化后产生 ATP 和 dATP。这类激酶在发育中的 T 淋巴细胞和 B 淋巴细胞内含量丰富，产生的高浓度的 dATP 可抑制核糖核苷酸还原酶，同时也抑制 S-腺苷同型半胱氨酸水解酶，使细胞内甲基化受阻，而引起与甲基化有关的许多物质代谢障碍，因此发生严重的免疫系统功能障碍。美国洛杉矶儿童医院将 ADA 基因转入产前诊断为此病的新生儿脐血 $CD34^+$ 细胞，回输后发现患者体内有产生 ADA 的骨髓细胞，外周血 1%～3%的 T 细胞含有外源 ADA 基因，实验性治疗有明显的疗效。还有报道用 ADA 基因转染 T 细胞的回输治疗后，血中 ADA 活性增加了 10 倍，患者生存 2 年以上。

二、肿瘤的基因治疗可在多个环节进行

基因治疗最早是针对一些单基因遗传病进行的，但是随着人类对其他疾病分子机制的

深入了解、对许多疾病相关基因的分离和功能的研究，人们逐渐将基因治疗的策略用于如恶性肿瘤、心血管疾病、糖尿病及艾滋病等，尤其是对恶性肿瘤的基因治疗寄予极大的希望。目前已被克隆的恶性肿瘤相关基因很多，动物模型比较成熟，病人及亲属易接受，所以，恶性肿瘤的基因治疗研究日趋活跃，并取得了显著的成果。

目前有关恶性肿瘤发病的分子机制的研究主要包括以下几个方面；第一，环境因素与组织细胞中原癌基因和抑癌基因的相互作用，这一对矛盾的失衡最终导致正常细胞的恶性转化。第二，肿瘤细胞通过某些作用机制逃脱了机体免疫系统的监视或破坏机体的免疫系统。在免疫监视机制中，细胞毒性 T 细胞（CTL）、自然杀伤细胞（NK）和巨噬细胞具有特别重要的作用。第三，细胞的异常增生和凋亡不足部分缘于细胞信号转导网络的异常，如生长因子、生长因子受体、蛋白激酶和转录因子等重要信号分子和信号转导分子的结构与功能异常。上述这些环节都可以导入某些基因进行干预，许多恶性肿瘤的基因治疗方案也正是以这些研究结果为基础而设计的。到目前为止世界各国已经批准开展进行的基因治疗方案中，70%以上是针对恶性肿瘤的。

（一）针对癌基因的治疗

许多恶性肿瘤的发生与发展与某种癌基因被激活后过表达有关，因此封闭某些癌基因或者抑制其过表达是肿瘤基因治疗中常用手段。抑制癌基因表达的方法有反义 RNA、核酶、RNA 干扰技术等。这些方法都是从 mRNA 水平抑制癌基因的表达。

（二）针对抑癌基因的基因治疗

研究表明，几乎一半的人类肿瘤均存在抑癌基因的失活，可见抑癌基因的失活与肿瘤的生长有着密切的关系。因此，将正常的抑癌基因导入肿瘤细胞中，可以补偿和修复突变或缺失的抑癌基因，达到抑制肿瘤的生长或逆转其表型。这是肿瘤基因治疗中的一种重要模式。

P53 基因是目前研究最广泛和深入的抑癌基因。其正常功能的丧失，最主要的方式是突变。在 P53 蛋白的 393 个氨基酸中有 280 个以上位点可发生突变。大量的体内外试验证实，引入 *P53* 基因可以抑制肿瘤细胞的生长，诱导其出现凋亡。2003 年 12 月我国深圳赛百诺基因技术有限公司推出带有野生型 *P53* 基因的腺病毒肿瘤基因治疗产品。这是世界上首次推出的商业化基因治疗药物。

（三）针对免疫系统的基因治疗

正常机体经常有很多细胞发生突变，但并不发生肿瘤，这是因为机体免疫系统通过细胞免疫机制能识别并杀伤突变的细胞。当肿瘤发生后，机体可通过免疫效应机制发挥抗肿瘤作用。机体对肿瘤的免疫主要以细胞免疫为主，针对免疫系统的基因治疗可以有效增强抗肿瘤作用。例如，将某些细胞因子的基因与载体重组后导入受体细胞内，可使其在细胞内部持续地产生细胞因子而达到激活和增强机体特异性抗肿瘤免疫反应的作用。

近年来，人们发现并克隆了多种人类肿瘤相关抗原基因，逐步了解了肿瘤抗原的识别和提呈及 T 细胞活化机制，并能成功地大量培养和扩增树突状细胞（dendritic cell，DC），从而使以 DC 为基础的抗肿瘤主动免疫治疗有了新的发展。DC 细胞是细胞膜表面高表达 MHC Ⅱ类分子的细胞，为人体最有效的抗原提呈细胞，能致敏和激活静止 T 淋巴细胞和 B 淋巴细胞。比较成熟的制备活化 DC 细胞的方法是采用抗原基因或细胞因子基因转染和修饰 DC 细胞。用不同形式的肿瘤抗原（蛋白或多肽抗原、肿瘤细胞溶解物、凋亡的肿瘤细胞）致敏的 DC 体内注射后可诱导特异性抗肿瘤免疫应答。大量荷瘤动物模型和临床试验证明，DC 免疫疗法具有确切的肿瘤治疗效果。DC 细胞作为肿瘤生物治疗和基因治疗的方案已获 FDA 批准并进入Ⅲ期临床。

（四）自杀基因治疗或酶药物前体疗法

自杀基因（suicide gene）指可引起细胞死亡的基因。例如将某些细菌、病毒和真菌中特有的药物敏感基因导入肿瘤细胞，并在肿瘤细胞特异性表达的基因调控区控制下，通过此基因编码的特异性酶类将原先对细胞无毒或毒性极低的药物前体在肿瘤细胞内代谢成有毒性的产物，以达到杀死肿瘤细胞的目的，也称药物敏感基因（drug sensitive gene）。这种方式可以减轻化疗药物的副作用。

（五）抗血管生成基因治疗

肿瘤的生长和存活依赖于血管为它提供氧气和营养物质。VEGF 是肿瘤诱导血管生成过程中一个主要的调节因子，它可选择性刺激内皮细胞分裂，并能增加微血管的通透性。通过阻断 VEGF 的翻译和转录过程可使血管的产生受到抑制。前述的反义 RNA 和小干扰 RNA 都可以用来抑制 VEGF 的合成。

三、心脑血管病的基因治疗

心脑血管病包括单基因控制和多基因控制的两大类，前者发病率较低，主要有家族性高胆固醇血症、遗传性扩张性心肌病、肥厚性心肌病、长 Q-T 综合征、遗传性心肌营养不良等。这些疾病是由某个基因突变引起，基因治疗的靶基因明确。相比之下与多基因有关的心脑血管病如高血压、动脉粥样硬化、心肌肥厚、心功能不全、再狭窄等是遗传因素和环境因素相互作用引起的，涉及多种基因结构和表达调控机制的改变，其发病的确切分子机制尚不十分了解。因此，多基因遗传的心血管病基因治疗是目前心血管病研究的难点。

利用过表达血管内皮生长因子 VEGF 促进血管生成是冠心病基因治疗方案之一，基因表达维持 3～4 周即可达到治疗目的。为此，可以采用携带有 VEGF 的 cDNA 的质粒直接注射。与病毒型载体相比，其安全性方面无疑具有很大优势，有利于临床研究的开展。此外，病毒载体在冠心病的基因治疗中仍占有较大的比重。

基因导入的途径即可以采用间接体内法（*ex vivo*），也可以直接在心脏内直接导入。前者的优点是在一个可控的体外环境中进行研究可得到最有效的基因转移，还可以避免载体的全身暴露及随后的免疫刺激。后者是相对简单，不存在供体血管、心肌细胞和严格细胞培养等方面的问题。

四、HIV 感染的基因治疗策略是抑制病毒复制或提高机体的免疫功能

艾滋病基因治疗研究主要在两方面取得较显著进展：一是通过基因修饰易感靶细胞，使之表达能特异性干扰 HIV-1 复制周期中一个或多个步骤的核酸序列或蛋白质，直接干扰病毒复制；二是提高机体的免疫能力、增强对 HIV-1 感染的免疫反应或采用靶向性毒素以杀伤和减少病毒感染细胞。

（一）直接干扰 HIV 病毒复制

1. 以 RNA 为基础的治疗策略　主要包括反义 RNA、核酶、RNA 诱饵及 RNA 干扰等。可选择的靶基因有 *tat*、*rev*、*env*、*gag* 及 5′端的引物结合序列等。

2. 以蛋白为基础的治疗策略　通过突变 HIV 基因组编码的某些蛋白质，将其导入 HIV 感染的靶细胞，产生负显性突变表型，如 *gag*、*rev*、*tat*、*env* 的某些表型的改变，以干扰病毒的反转录、病毒蛋白的合成和释放。研究较多的病毒蛋白突变体是 RevM10。

采用定点突变技术，使野生型 Rev 蛋白第 78、79 位的亮氨酸、谷氨酸分别被门冬氨酸、亮氨酸取代。这种突变蛋白无致病性，但可以与野生型蛋白竞争结合位点，对 HIV-1 复制产生剂量依赖性的抑制作用。

（二）增强机体的免疫功能

特异性 $CD8^+$ T 细胞免疫反应与 HIV-1 复制的抑制密切相关。增强针对 HIV-1 抗原的免疫反应，激活强有力的细胞毒性 T 淋巴细胞反应可以帮助机体免疫系统清除和破坏病毒感染细胞。

（三）造血干细胞用于抗 HIV-1 基因治疗

如果体内新生的 $CD4^+$ T 淋巴细胞和单核巨噬细胞都来源于含有抗 HIV-1 基因的造血干细胞，这些细胞则可能免于 HIV-1 感染的破坏。因此，将抗 HIV-1 基因（如 HIV 辅受体的配体）导入造血干细胞，再回输到 HIV 感染病人体内，可能达到永久抑制 HIV-1 复制。

五、开展基因治疗必须遵守国家法规

经过 20 多年的努力，科学家们在基因治疗领域取得了很大的进步，获得了一些成功，但是仍然存在一些亟待解决的问题。这些问题包括：①缺乏高效、靶向性的基因转移系统；②对于多种疾病的相关基因认识有限，因而缺乏切实有效的治疗靶基因；③目前真核生物基因表达调控机制尚未完全阐明，因此对治疗基因的表达还无法做到精确调控，也无法保证其安全性；④缺乏准确的疗效评价。目前的基因治疗临床试验中，限于伦理问题，多选择常规治疗失败或晚期肿瘤患者，尚难以客观地评价治疗效果。

基因治疗的安全性是保证这种治疗方法大规模应用的关键，因为在基因治疗的发展过程中，曾出现过几起严重的不良反应，使很多人对这种方法心存芥蒂。这里的安全性不是指药物的副作用，而是基因治疗是否会给病人带来致瘤、致畸和遗传性的重大问题。争议的焦点主要集中在用病毒作载体进行目的基因的转移具有潜在的危险性，如反转录病毒作载体可能随机整合到宿主染色体基因组中，有致瘤、致畸的可能性。法国就曾经发生过重症联合免疫缺陷病儿童接受利用反转录病毒载体的基因治疗后发生白血病的病例。

将基因治疗方案用于人体必须经过严格的审批程序，需要专门机构的审批与监督。美国是最早开展基因治疗的国家，每个用于临床基因治疗的方案需经过几个机构的审查，包括地方伦理小组、地方生物安全小组，国立卫生研究院重组 DNA 顾问委员会人类基因治疗分委员会、食品药品管理局等。

在我国，任何基因治疗方案都要经国家食品药品监督管理局审批。鉴于基因治疗重要的科学意义及巨大的应用价值，为促进和指导我国基因治疗研究、开发和产业化的正常发展，1993 年 5 月我国卫生部公布《人的体细胞治疗及基因治疗临床研究质控要点》；1996 年，国家食品药品监督管理局对上述质控要点作了进一步修订，并于 1999 年 6 月颁布了《人基因治疗申报临床试验指导原则》。在这些法律法规中详细地规定了基因治疗所用生物制剂的研制、生产工艺、制剂的质量控制、临床实验和临床疗效评价的各个环节中应该遵守的原则。所有基因治疗的临床使用必须严格遵照这些法律法规。

（葛银林）

第二十四章　分子生物学与药物发现

关键知识点

- 传统的药物发现有两种主要模式：一是随机发现药物，二是定向筛选发现药物。现代药物发现策略则是通过高新技术的应用，以疾病机制的认识为依据、以药靶结构为基础进行合理的药物设计，以及利用基因和遗传学的发现寻找有价值新靶，进行逆向药理学研究。未来药物发现模式则是由线性过程向综合、集成、启发、探索性的过程转变，包含有一系列的研究反馈环路的环形反馈模式。
- 运用分子生物学技术可揭示疾病发生、发展的分子机制，为药物发现提供了新靶点；基因组学的研究也提供了新的药物靶点。常用于药物靶点发现和验证的分子生物学技术主要有：全长 cDNA 文库构建、差异基因表达解析、转基因与基因打靶技术、反义寡核苷酸技术、RNAi 技术、蛋白质组学技术等。用于药物筛选的主要新技术有：高通量筛选技术、生物芯片技术、基因工程制备筛选模型（重组受体）、转基因动物模型、组合化学技术及其他技术。
- 随着药物基因组学和药物蛋白质组学研究的兴起，后续转录组、代谢组、化学基因组等“组学”计划的实施，生物信息学与系统生物学也取得突飞猛进的发展，这些学科对药物发现过程将产生深远的影响，结合计算机网络、系统信息整合，为药物发现提供重要的理论基础，并推动药物发现进入了一个革命性变化的新时代。

关键词： 药物；药物发现；分子生物学；靶点；合理药物设计；逆向药理学；药物基因组学；药物蛋白质组学

医药科学研究是人类进步过程中极其重要的领域，对维护和促进人类健康发挥了巨大的作用。传统的药物发现过程很大程度上借助于机遇及定向筛选过程，而现代的分子生物学、遗传学提供了包括 DNA 序列和基因种类信息在内的大量生物学信息，不仅使我们对药物作用靶点与发病机制之间的关系有了足够的认识，也使我们从更深的层面——基因分子水平上去认识生物界所发生的一切，使药物的发现和开发进入一个新的历史阶段，即以基因为基础的药物发现和开发阶段，为药物研究开辟了新的道路。

本章将从药物发现模式的变革、分子生物学技术在药物发现中的应用及现代生物学与药物发现三个方面阐述分子生物学与药物发现。

第一节　药物发现模式的变革

药物（drug）是指可查明或改善机体的生理功能或病理状态，对用药者有益，达到预防、诊断、治疗疾病和计划生育目的的物质。药物发现（drug discovery）广义的定义是药物研究和开发（research and development，R&D）过程，包括有：某种疾病和治疗靶点确定的基础研究和可行性分析；先导化合物体内外检测的生物模型和方法学的建立；药理学、药物代谢动力学和安全性研究；制剂学；专利申请以及人体Ⅰ、Ⅱ、Ⅲ期临床研究和上市销售。狭义上指药物设计（drug design），即基础研究和可行性分析设计的先导化合物发现过程。

一、传统的药物发现模式

药物发现和研究的最终目的是用于临床治疗疾病，因此，在药物发现过程中，对药物的最基本要求就是安全有效。无论药物作用的方式或机制如何，只要能够解除病痛，治疗疾病，就是有效的药物。目前应用的药物基本上是以改善症状作为基本评价标准，因此对药物改善症状表现的作用研究最为充分；然后是对组织器官功能影响的研究，而对分子基因水平变化的研究，目前仍然处于尝试和探索阶段。只有明确了基因与疾病之间的关系，才能真正地使药物作用于基因水平，改变基因或基因的表达，治疗疾病。

分析药物发现的历史，可以将药物发现的过程归纳为两种主要的基本形式：一是随机发现药物，包括治疗作用的随机发现和药物来源的随机发现，在这种形式下，后者是随机发现的主要内容；二是定向筛选发现的药物，也包括对治疗特定疾病药物的筛选和特定药物来源物质的筛选，在这种方式中，前者是主要内容。

传统的药物发现模式是从药用植物中分离的有效成分；或对这些成分的结构进行改造并人工合成的化合物；或是微生物发酵产物；运用动物模型进行定向药理筛选或随机筛选，发现有活性的先导化合物，继而进行优化研究，再通过进一步药理学筛选，确定候选药物，进行临床前试验和评估，及临床试验和评估，从而产生药物，上市销售。

二、现代药物发现策略

新技术的应用，如分子生物学、高通量筛选技术、计算机辅助药物设计技术、组合化学和组合生物技术等，必然会对传统的药物发现模式产生冲击。人类的药物发现经历了从自然界发现药物，如神农尝百草，随机筛选发现药物，到以机制为基础和以药靶结构为基础的新药发现和开发，该过程也是一个从盲目发现药物到理性设计和发现药物的历程。

1. 合理药物设计　发现先导化合物是药物发现的关键一步。目前要发现先导物可通过两个途径。一是从众多的化合物中挑选出有活性的先导物，称为筛选，大容量的化合物库是筛选研究的物质基础；二是以生命科学为基础，根据机体内生命活性物质以及与疾病相关的药物作用靶点设计先导物，即理性的药物设计，这也是更富挑战性的捷径，分子生物学在其中起重要作用。

合理药物设计（rational drug design）是依据药物发现过程中基础研究所揭示的药物作用靶点，即广义上的受体，如酶、受体、离子通道等，再参考其内源性配基或天然药物的化学结构特征，寻找和设计合理的药物分子，以发现既选择性作用于靶点又具有药理活性的先导物，或根据靶点的三维结构直接设计活性配基的过程。

合理药物设计可分为三种类型：

（1）受体结构已知，即通过结构测定技术得到了受体或受体-配体复合物的三维结构，这样可在受体活性部位直接进行药物设计；

（2）受体的一级序列已知，但尚未阐明其三维结构，可根据同源蛋白质建模等方法，预测其三维结构，然后在根据预测出来的模型结构进行药物设计；

（3）受体结构尚未阐明，只有一些配体的结构知识，则根据其配体推导出药效基团模型，甚至提出一个假想受体，再间接进行药物设计。

总之，受体即标靶的结构知识掌握得越多，设计出来的配体就越有希望成为药物。合理药物设计是以生命科学和现代仪器及技术为基础的靶向信号转导分子和围绕药物作用靶点设计药物的一种方法和技术，具有指导性的基本原理和概念，应用前景宽广。

2. 逆向药理学　以基因为基础药物发现研究的主要过程为：基因组→新靶→筛选→先导物→药物。人群中遗传变异会影响药物的代谢或产生不必要的毒副作用，这对药物发现是非常重要的。但对药物发现更直接的影响是药物作用新靶的发现，这些新靶一般是由新发现的基因所编码的。靶与疾病病理学有关的特异性和选择性是衡量它能否作为药物靶的重要依据，利用基因和遗传学的发现是寻找有价值新靶的关键。

传统的药理学研究模式沿袭的基本途径是：底物（药物）→受体→基因，即发现特异性作用于某一类受体或受体亚型的底物或药物从而确定受体的存在，然后通过各种方法分离得到相关受体，再设计寡核苷酸探针并从 cDNA 库中分离得到相关基因，最后确定各种受体的超基因家族。这一过程的反向模式即从基因→受体→药物，即以基因为基础的药物发现称为逆向药理学（reverse pharmacology）。首先从各种受体超基因家族中分离得到第一代基因，往往会提示基因族中伴有大量结构相似性，其结果是同族受体氨基酸序列蛋白结构的相似性。因此对 DNA 序列可用各种生物技术在基因水平上进行分离，目前常用的方法有两种：一是用一系列兼并探针筛选分析 cDNA 或 DNA 库；二是在已知受体基因序列的基础上，用兼并引物通过 PCR 反应进行 cDNAs 的放大。这两种方法均可对哺乳动物细胞表达的未知受体的基因进行克隆。G 蛋白偶联受体超基因家族的某些新的成员就是用 PCR 技术完成了受体基因克隆，如腺苷受体、大麻碱受体，以及一些尚未明确配体的类固醇受体。这一过程不仅证明了各种受体的存在，更重要的是给药物发现提供了新的生物学基础。

随着一些分子生物学新技术如：DNA 突变检测技术、基因敲除或沉默技术、高通量 DNA 测序技术、基因分型技术、基因克隆和表达技术等技术的不断改进和完善，从基因到药物的过程将会变得更加方便和快捷。

三、未来药物发现模式

传统及现代的药物发现模式是一个线性过程，即：靶点的发现和验证→蛋白质检测，高通量筛选→偶尔发现先导物→临床前试验→Ⅰ期临床→Ⅱ期临床→Ⅲ期临床→FDA 或国家药品监督管理局认证→Ⅳ期临床。该线性过程没有或少有反馈信息，或者说是下游的

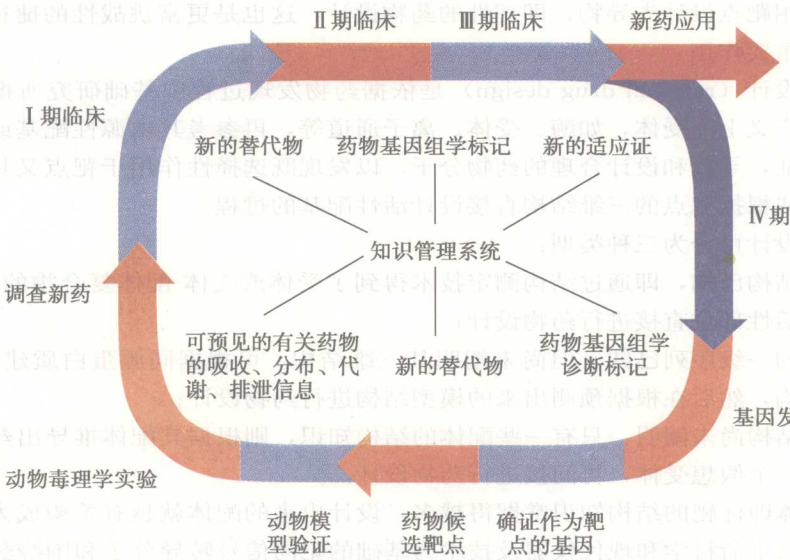


图 24-1　未来的药物发现模式

结果对上游的化合物没有发挥明显的改善和改进作用。

有关疾病分子基础研究的进一步深入和基因组学的迅速发展将改变这一模式，个体化用药策略的应用将迫使药物发现过程由线性过程向综合、集成、启发、探索性的过程转变。这一新进展将包含有一系列的研究反馈环，具体图示如图 24-1。在药物发现的早期阶段，如药物靶点的选择与验证、小分子筛选、化学处理及化合物的临床前评估等，将与临床研究的后期阶段紧密相连。在不同阶段所获得的分子、药理学及患者的临床资料将被整合成一个“知识管理系统”(knowledge management system)，有利于促进分子疾病的理性药物设计。知识管理系统中将存储有后续过程及产品改进方面的任何有用信息。在药物发现的每个阶段出现的基因组学信息和标记都将作为药物开发上游过程和下游过程中的工具，从而产生更好的药物制剂和个体化药物。

第二节　分子生物学技术在药物发现中的应用

一、药物靶点

药物作用的靶点（target）即广义的受体，是生物体的细胞膜上或细胞内的一种特异性大分子结构。药物和信息分子能与有关受体大分子的关键部位（受点）特异性结合，生成可逆性复合物，并进一步启动功能性变化，如开启细胞膜上的离子通道，或激活特殊的酶，从而导致生理药理变化。

药物作用的靶点有多种，如受体、酶、离子通道、抗原、核酸、糖类大分子及脂质等，各有其功效。一种疾病可有多个药物作用的靶点，例如抑制胆碱酯酶、激动 M 胆碱受体、激动 N 胆碱受体、拮抗腺苷受体、拮抗 5-羟色胺受体、激动促甲状腺释放激素受体、调节钙通道、调节钾通道等均可治疗阿尔茨海默病，这样阿尔茨海默病就存在上述多个治疗靶点；另外，某一靶点也可是多种疾病共同的治疗靶点，例如环氧化酶-2（Cox-2）是花生四烯酸转化为前列腺素途径中的一种关键酶，与炎症、阿尔茨海默病、癌症、动脉粥样硬化等密切相关，从而成为药物作用的新靶点。

运用先进技术（量子机制、药物-受体的计算机模拟、单晶 X 线衍射、多维磁共振谱、基因克隆与表达等）的相互作用和结构特征为枢纽进行生命科学的基础研究，发现疾病治疗的多种靶点以及相关的新化学实体（new chemical entity，NCE），是药物发现阶段中最重要和最有意义的，也是最富挑战性和最有风险的首要一步。20 世纪新药研究集中于细胞膜上的酶和受体靶，到 21 世纪已扩大到核酸和糖类等，进行细胞和基因的修饰和调控，研究重点向细胞核内转移。

二、分子生物学研究提供新的药物靶点

现代分子生物学技术的发展已能确定并克隆许多调节正常生理功能的蛋白质，并阐明它们在疾病状态下的作用，解释信号转导失控，揭示引起许多疾病的分子机制。如何发现外源性药物进行调控，目前有两条重要途径：一是以信号分子作为药物设计的靶点，或以内源性活性物质为先导物进行研究；二是对疾病机制（即信号转导途径的差异）进行基础性研究以求发现药物作用的靶点，其基础是内源性活性物质调节信号转导作用的分子机制。

（一）基于机制的药物设计

根据疾病的发病原因和药物防治疾病的机制，针对其关键环节和限制性步骤，同时考

虑药物在体内的转运与代谢来设计化学药物，称为基于机制的药物设计（mechanism-based drug design，MBDD）。以艾滋病为例，说明如何围绕疾病机制，寻找治疗靶点。艾滋病（AIDS）是由人类免疫缺陷病毒（HIV）感染所致。在HIV增殖周期过程中，病毒基因组核酸的复制和蛋白质的合成是最关键的步骤，需要有某些特异性酶的参与，所以反转录过程是HIV最重要的特征之一。HIV反转录酶是HIV从mRNA转录为DNA的过程中起主导作用的酶，因此也就成为研究抗艾滋病药物最佳的靶点之一。另外，HIV蛋白酶是HIV产生的，有两种HIV蛋白产物是裂解成熟蛋白的前体，经HIV蛋白酶的催化，所释放出来的蛋白对病毒的复制起决定性的作用。如果这些蛋白不从前蛋白中裂解出来，HIV就无感染性，因而HIV蛋白酶也就成了又一个重要的靶点。目前已经研究出多种HIV反转录酶抑制剂（齐多夫定 Zidovudine、扎西他宾 Zalcitabine、去羟肌苷 Didanosine、奈韦拉平 Navirapine 等）和HIV蛋白酶抑制剂（沙喹纳韦 Saquinavir、英地纳韦 Indinavir、奈非纳韦 Nelfinavir 等）用于抗艾滋病。对艾滋病发病机制的进一步研究将伴随着新的药物靶点和新的药物出现。

（二）基因组学的研究在新靶发现中的作用

人类基因组计划的启动和完成，增加了我们对靶点的选择范围。2001年2月，国际人类基因组科研小组和美国塞莱拉公司同时公布，人类整个基因组的基因数量在2.5万左右，但90%以上的基因尚不知功能。在这些基因序列中，出现了大约5 000个药物作用靶点。到目前为止，已经有近千个靶点类蛋白被克隆纯化，包括大约750个G蛋白偶联受体、100多个配体门控性离子通道、60多个核受体、50多个细胞因子和大约20个具有重吸收和转运功能的蛋白质。如何将那些未知功能的基因转化为有治疗前景的药物靶点也就尤为重要。现代分子生物学技术可为我们提供有效的分子靶点，除了蛋白质产物以外，基因本身也可以作为药物的靶点。

三、分子生物学技术在药物靶点的确认与验证中的应用

现代药物研究中，新靶点的建立往往是新药创新的前提和保障，因此对新型靶点进行选择和验证便成为非常重要的工作。分子靶点的选择和验证需要许多分子生物学技术作支撑。常用于发现和验证药物靶点的分子生物学技术主要有：

1. 全长cDNA文库构建　作为人体设计图的情报，全部在基因组上记载，在测定基因组序列之后，解析基因的功能，全长cDNA文库（见第十九章）的构建是必不可少的。例如美国Incyte数据库，从正常人的脏器和从癌症患者身上摘取的脏器等制作的各种cDNA文库，从目的基因的表达量在各文库的出现频率可以粗略地进行推测，文库中表达量高的基因有可能成为抗癌药物的靶点。在构建cDNA文库的基础上，灵活应用生物信息学的各种软件，可以非常迅速地进行同系解析，寻找与药物开发有关的靶分子。

2. 差异基因表达解析　基因表达谱的研究，是以探讨与疾病有关的基因和阐明药物的作用机制为目的（见第二十章）。在疾病组织，基因的表达水平往往与正常组织不同，可能表现为基因表达上调、下调和基因的完全缺失等，因此对差异基因表达进行定位是鉴定某个分子靶点的重要步骤。另外，靶点在某个组织上呈现特异性表达，可能是药物设计中限制药物副作用的基础，因此研究疾病中差异基因的表达在药物开发中作用极为重要。有高通量基因表达分析平台之称的DNA芯片/DNA微阵列技术，可在一个样本上同时研究成百上千个差异基因的表达，使差异基因解析过程更迅捷，并进一步加快了药物开发的进程。

3. 转基因与基因打靶技术　通过转基因技术（见第十九章）制备转基因动物模型，

可获得关于可能作用靶点的重要的在体功能学信息。基因打靶包括基因敲入和敲除技术，无论是基因敲入而过表达还是基因敲除无相应蛋白的表达，在药物发现过程中的靶点的确认都是相当有用的，因为某个药物靶点选择以后，通过基因敲入或敲除试验以检测该基因过表达或缺失以后是否就出现或丢失了某种功能，从而对该靶点进行确证，这样还可排除一些因为环境因素而非真正的基因变异所致的假阳性结果。一般以基因敲除法制作动物模型较基因敲入法更为多见。另外，转基因动物还可用于制备重组药物。

4. 反义寡核苷酸技术　在完整细胞中，反义寡核苷酸能特异性抑制某一目的蛋白的合成，这在药物靶点的鉴定和药物设计方面起着重要的作用（见第二十二章），并且有些反义寡核苷酸本身就可作为药物成功地应用于临床试验，具有较好的应用前景。

5. RNAi 技术　RNAi 是一种高通量破坏基因功能的方法（见第十九章），通过特异性的 dsRNA 敲除靶基因，而并不产生突变。因此，RNAi 在药物发现过程中可以对靶基因进行高通量的功能分析，有助于鉴定药物作用的生化模型和基因编码产物与特异化合物的相互作用，或对特异化合物的反应。RNAi 技术将在药物开发及疾病治疗方面发挥越来越大的作用。

6. 蛋白质组学　蛋白质组学是应用二维凝胶电泳、质谱测序、酵母双杂交系统、适体法等新技术（见第七章），从整体水平上研究蛋白质功能，可以获得全套基因的表达产物信息。比较正常和疾病状态的蛋白质表达，对与疾病有关的蛋白质进行定性、定量化分析，可发现药物的新靶点。从理论上了解人类基因组中的每一种蛋白质的确切结构，将促进理性的药物设计。目前只有大约 1%的人类蛋白质的结构已被确定，随着人类蛋白组学的发展，必将给药物研发带来新的契机。

四、分子生物学技术在药物筛选中的应用

从众多的化合物中挑选出有生物活性的先导物的过程，称为筛选，也就是说必须经过筛选才能确定先导物。传统的新化合物筛选是以分子功能分析和改进为主要手段，通过大量化合物合成与结合实验来确定可能有效的新药，这是一低效率的过程。近年来，根据组合化学理论设计的各种化合物文库开始用于药物筛选，其中可放大的生物文库的建立和应用，以及高通量自动化药物筛选技术的出现，意味着大批量的化合物可以在很短时间内快速地进行筛选。

组合化学（combinatorial chemistry）起始于 20 世纪后期，基本原理是采用适当的化学方法，在特定的分子母核上加入不同的基团，在同样条件下，产生大量的新化合物。这种方法在化合物的结构改造和优化方面表现出强大的优势。但是，由于该方法是基于母核结构的改造，因此产生的大量化合物在结构多样性方面尚有极大的不足。但是，由于这项技术可以产生人工的“天然产物”，因而受到研究人员的重视。采用组合化学的方法，可以产生大量化合物，每周约可达到上千个化合物，扩大了药物发现的范围。

高通量筛选（high throughput screening，HTS）是应用计算机控制的先进的实验方法和技术，对大量化合物进行微量样品的自动化测定，是 20 世纪后期起发展起来的新药发现新技术。高通量药物筛选主要依赖于含有大量化合物的样品库、计算机控制的自动操作系统和微量灵敏的生物反应及检测系统。高通量药物筛选是采用体外的实验方法（主要是在分子和细胞水平）评价药物活性的自动化大规模筛选过程。因此，由此筛选的结果可以明确阐明化合物生物活性的分子机制，但在评价其整体动物药理反应方面，还受到多种因素的影响，存在较大的差异，需要系统的后续工作进行评价。

用于药物筛选的主要分子生物学技术有：

1. 生物芯片技术　应用芯片技术（见第二十一章）对生物文库进行药物筛选，其基本原理是在芯片或合成珠等固相表面上进行原位文库合成和筛选。生物芯片技术使得药物筛选、靶基因鉴别和新药测试的速度大大提高，甚至能在几天内确诊某种药物是否会导致遗传突变或造成其他危险，以及通过芯片技术对病人群体进行分类，从而加快新药审批程序。芯片技术还可用于中药有效成分的筛选。另外，芯片技术的微量化和分析自动化使得被检样品的需要量极微，可测的样品的数目则成千万倍地提高，以至药物筛选的成本大大降低。

2. 基因工程制备筛选模型——重组受体　在新药研究开发中日益广泛使用的各种酶和受体筛选模型所需的靶酶和受体往往来自动物体内，因而数量有限且不利于大量筛选。应用基因重组技术制备克隆受体。把它们用于筛选模型，可大大提高对人类的针对性。其优点很多：①受体是人的而不是动物的，所以用重组受体所得的试验结果与在人体试验的结果直接相关；②可大量制备受体，并能得到只在特定组织中存在的用传统制备方法难以获得的受体；③用克隆方法可得到纯的受体；④重组受体在细胞内或细胞膜上，包含了G蛋白等信号转导系统，更接近人体的自然状态；⑤重组受体更经济等。

3. 转基因动物模型　转基因动物是（见第十九章）当今分子药理学研究的重要手段，是作为疾病模型用于药物筛选的一种重要工具。利用转基因动物可建立敏感动物品系及与人类相同疾病的动物模型用于药物筛选，避免了传统的动物模型与人类某种症状相似的疾病在病因、发病机制方面不尽相同的缺点，其结果准确、经济，试验次数少，大大缩短试验时间，现已成为人们试图进行药物“快速筛选”的一种手段。

其他重要技术，如三维核磁共振（3D-NMR）技术、X射线晶体学、扫描隧道显微镜技术（STM）及计算机分子模型技术（CAMM）等，均在研究分子的三维结构方面起重要作用，有助于理性药物设计，促进药物发现过程。

第三节　现代生物学与药物发现

自20世纪80年代以来，随着人类基因组计划的完成以及后续功能基因组、化学基因组、比较基因组、转录组、蛋白质组以及代谢组等“组学”（omics）计划的实施，生物信息学与系统生物学也取得突飞猛进的发展，这些现代生物学分别从基因（DNA）、蛋白质、RNA（mRNA）、代谢产物等多个层面对药物发现过程产生深远的影响，并结合计算机网络、系统信息整合，推动药物发现进入了一个革命性变化的新时代。

一、药物基因组学

（一）药物基因组学的概念

药物基因组学（pharmacogenomics）是一门研究遗传变异对药物效能和毒性的影响，或者说是研究患者的遗传组成如何决定对药物的反应的科学，是出现在后基因组时代的制药工业中的一门极有前景的科学。其大致含义包括有：药物效应的基因型预测和基因组学在医药工业上的应用，在分子水平阐述和证明药物疗效、药物作用的靶位、模式和毒、副作用。药物基因组学以提高药物的疗效及安全性为目标，研究影响药物吸收、转运、代谢、消除等个体差异的基因特性，以及基因变异所致的不同病人对药物的不同反应，并以此为平台开发新药，指导合理用药，提高药物作用的有效性、安全性和经济性。

（二）药物基因组学的研究内容

药物基因组学区别于一般意义上的基因组学，它不是以发现人体基因组基因为主要目

的，而是应用已知的基因组学理论改善对患者的治疗。药物基因组学不能简单地看成是预测易患病的体质，或者提供统计学资料用于预测发病危险性的试验，而是有关基因对药物作用本身产生的影响及与基因变异的关系，这些变异正是不同个体产生不同药物效应的根本原因。因此，基因多态性是药物基因组学的基础和重要研究内容。

药物基因组学所研究的基因多态性主要包括药物代谢酶、药物转运蛋白、药物作用靶点等基因多态性。药物代谢酶多态性由同一基因位点上具有多个等位基因引起，其多态性决定表型多态性和药物代谢酶的活性，并呈显著的基因剂量-效应关系，从而造成不同个体间药物代谢反应的差异，是产生药物毒副作用、降低或丧失药效的主要原因之一；转运蛋白在药物的吸收、排泄、分布、转运等方面起重要作用，其变异对药物吸收和消除具有重要意义；大多数药物与其特异性靶蛋白相互作用产生效应，药物作用靶点的基因多态性使靶蛋白对特定药物有不同亲和力，导致药物疗效的不同。

药物基因组学研究的主要策略是选择药物起效、活化、排泄等过程相关的候选基因进行研究，鉴定基因序列的变异，在分子水平研究其在药物作用中的意义，也可用统计学原理分析基因突变与药效的关系。药物基因组学的发展需要高效的基因变异检测方法等技术支持，基因变异检测技术检测药物相关基因的变异情况，为鉴定遗传变异对药物作用的影响提供了前提条件，并可能对药物基因组学产生重大影响。

药物基因组学将广泛应用遗传学和基因组学信息来预测患病人群对药物的反应，从而指导临床试验和药物开发过程，还将被用于临床患者的选择和排除，并且提供区别的标准。新的基因组学技术，如基因变异检测技术、DNA 芯片技术、SNP 研究的高通量技术、药物作用显示技术、生物统计分析技术、基因分型研究技术及蛋白质组学技术等，为药物基因组学的进一步研究和应用提供技术支持并开辟了道路。

二、药物蛋白质组学

药物蛋白质组学（pharmacoproteomics）是以蛋白质分析为基础的将患者进行分类的研究。其重要研究内容，在临床前包括发现所有可能的药物作用靶点，以及针对这些靶点的全部可能的化合物，药物作用机制和毒理学研究，在临床研究方面包括疾病特异性蛋白作为有效患者选择的依据和临床试验的标志，或应用类似于药物遗传学的方法，按照蛋白质谱来分类患者，并预测药物作用疗效。

蛋白质水平的分析不仅为生物分子体系提供了最有效的实时分析模式，而且也能获得在 DNA 和 RNA 水平上不易获得的信息。例如，对肝癌患者的最近研究表明，甲胎球蛋白（AFP）的糖基化变化是比甲胎球蛋白更为灵敏的指标。在这种情况下，正常和异常的糖基化甲胎球蛋白在氨基酸序列上则没有差异，应用基因水平的分析则不可能发现这种差异。二维凝胶电泳和质谱的结合与精致的生物信息学建立了一套强有力的、通用的方法，用于蛋白质的分布作图和确定蛋白质的结构特征，使得药物蛋白质组学至少具有与遗传学水平上相同的作用。

三、其他现代生物学方法与药物发现

随着药物基因组学和药物蛋白质组学研究的兴起，转录组学、代谢组学、化学基因组学等“组学”计划的实施，生物信息学与系统生物学也取得突飞猛进的发展，为新药开发提供重要的理论基础，并成为药物发展的根本动力。

1. 转录组学　转录组学（transcriptome）研究可以精确反映基因表达的时空性（见第

七章），以及机体组织细胞对内/外环境变化的反应性和适应性，提示药物潜在的作用靶点及其发挥作用的分子机制。与基因组学相比，转录组学方法能更直接地得到功能基因的信息，而且比蛋白质组学更稳定，范围更广阔，是与基因组学相融合的一种发现功能基因的路径。转录组学研究在药物发现过程中也具有重要的应用价值，包括药物靶点的探索与发现、指导新药设计与合成、新药筛选等。

2. 代谢组学　代谢组学（metabonomics）（见第七章）是后基因时代的一门新兴的独立学科，它与药物的药效和毒性筛选以及评价研究、安全性评价、作用机制研究和合理治疗用药密切相关。代谢组学技术可以广泛地应用于新药研制从早期发现到临床开发的全过程。代谢组学研究“代谢指纹图谱”，它不仅研究药物本身的代谢变化，还要研究药物引起的内源性代谢物的变化，两者结合系统全面地反映体内生物化学过程和状态的变化，进而阐明药物活性部位与药效的关系、药效基团作用于病变靶点的本质过程。通过代谢组学技术对药物进行研究，可以在新药的开发中很好的监控药物的吸收、分布、代谢、排泄和对药物的毒性进行预测，从而能大大缩短药物发现和开发的周期，并且有助于人们在基因水平上认清药物作用体内后，内源性物质变化和药物作用的关联。

3. 化学基因组学　化学基因组学（chemogenomics）是在化学配基的合成、筛选和确认的精密技术快速发展的背景下产生的，整合了组合化学、基因组学、蛋白质组学、分子生物学、药物学等领域的相关技术，采用具有生物活性的化学小分子配体作为探针，研究与人类疾病密切相关的基因、蛋白质的生物功能，同时为新药开发提供具有高亲和性的药物先导化合物。化学基因组学，作为后基因组时代的新技术，是联系基因组和新药研究的桥梁和纽带。它使用对确定的靶标蛋白高度专一的小分子化合物来进行基因功能分析和发现新的药物先导化合物。对药物发现的化学基因组学途径与传统途径进行比较（图 24-2），

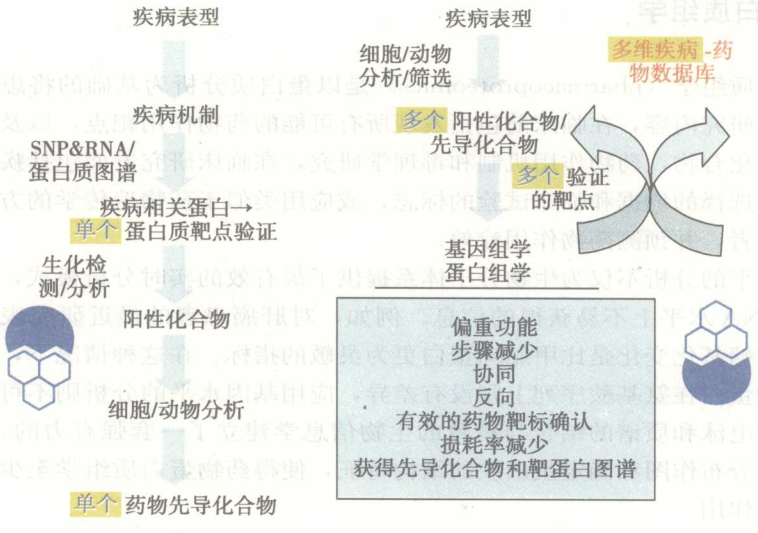


图 24-2　药物发现传统途径与化学基因组学途径比较

A. 传统药物发现途径；B. 化学基因组学途径

与传统药物发现途径相比，化学基因组学途径的优势主要表现为：①基于客观的以细胞为基础的分析所构建的疾病表型和网络，进行功能性筛选，可提高药物发现的效率；②与逐步的传统药物发现途径相比，步骤减少；③药物发现过程中有更多的协同与相互作用；④逆向/循环的药物发现途径；⑤通过功能学筛选可获得更多的有效的确认靶标；⑥可发现并绘制出多个药物/治疗靶点；⑦建立多维的药物-疾病资料库

二者在多个方面有所不同。化学基因组学在小分子药物研究中有独特的优势，促进了小分子药物的开发进程。

4. 生物信息学　生物信息学是随着人类基因组计划的启动而兴起的一门综合生物学、计算机科学以及应用数学等学科的交叉学科，研究内容主要有建立、贮存并管理大量的生物学信息库，包括基因组序列、基因多态性、基因表达调控、蛋白质结构与功能、特征性代谢产物谱、疾病相关基因和/或蛋白、生物标志物信息库等；开发计算机算法和统计学方法，分析确定数据库中大量数据相关性；应用已知的生物学信息预测或分析生物大分子或小分子化合物的结构与功能。生物信息学可应用于药物发现的全过程，包括药物分子设计、药物靶点的发现与确认、药物筛选以及药物临床前评价等。随着基因组研究的蓬勃发展，可利用的数据和信息日益庞大，生物信息学可以通过快速地分析、选择，发现和确定新的药物靶点，并通过计算机构建各种模型，方便、快捷地验证各种设想，指导生物活性筛选，从而设计或发现更为安全、高效的药物。

5. 系统生物学　系统生物学（system biology）不同于以往的实验生物学，不仅关注个别的基因和蛋白质，还要研究所有的基因、蛋白质等组分间的所有相互关系。同时，通过整合各组成成分的信息，包括从生物体内各分子的鉴定及其相互作用的研究到体内生命信号途径和网络的研究，以图画或数学方法建立能描述系统结构和行为的模型。系统生物学的技术平台为组学，即基因组学、转录组学、蛋白质组学、代谢组学、相互作用组学、表型组学和计算机生物学等。系统生物学将成为药物发现的新的技术平台，以精确的模型研究生理、病理现象，将原来孤立研究的生命过程联系在一起，实现网络理解。

现代“组学”研究当前还处在一个初期发展阶段，相关的技术手段及其配套应用也不够成熟。然而，随着这些学科研究的迅速兴起，人们将在基因水平和蛋白质水平上重新认识诸如生长、发育和代谢调控等生命活动的规律，为研究重大疾病的机制、疾病诊断和防治与新药开发提供重要的理论基础，成为药物发现的根本动力。

（臧伟进　陈莉娜）

第二十五章　医药生物技术产品的研究开发与管理

基因操作技术、基因工程技术、基因治疗技术、生物芯片技术、基因组研究技术、生物信息技术等各种新兴的生物技术已被广泛地应用于医药产品的研究和开发。药物的研发过程一般包括实验室研究、中试研究、临床前研究、临床研究等几个阶段，在这一过程中需要遵循一整套相关的管理制度，以保证获得质量合格的产品，并满足生物安全的要求。

第一节　医药生物技术产品研究开发的一般过程

一种新的生物技术产品从实验室研究开始到国家药品监督管理部门批准并进行正式生产和销售，是一个漫长的过程，一般要经过实验室研究、中试研究、临床前研究、临床研究等阶段，最后得到正式批准后才可进行大规模的生产并进行销售。这里对这个过程各个阶段的研究内容以及应该遵守的国家相关规定进行简要介绍。

一、实验室研究

生物技术产品的实验室研究阶段是通过对特定基因和蛋白质结构、功能的研究，并进行小量试制，探索其作为药物和诊断试剂的可能性。这一阶段的研究根据具体情况不同，需要的时间差异较大，少则 1～2 年，多则 10～20 年。在此研究阶段一般包括以下的研究内容：

1. 医药生物技术新产品的发现和筛选　医药生物技术新产品的发现一般需要经过大量的基础研究过程。研究者采用本书前面章节所述的各种技术和手段对研究对象进行大量的研究，发现某个靶位（基因或蛋白质等）具有可能的治疗或诊断意见，并对其研究开发的前景进行初步的评估，决定其是否作为药物或诊断试剂进行开发。一般说来，这些工作进行得越扎实，开发的风险性可能越小。

2. 实验室规模的小量试制　实验室规模的小量试制（小试）是对目标产品进行初步的、小规模的制备，用于其生物学活性或功能的初步研究，并确定以后大致的生产工艺。对于活性多肽和蛋白质以及基因治疗制剂需要采用基因工程的方法进行生产，核酸类药物常采用核酸合成仪进行人工合成，而单克隆抗体则采用杂交瘤技术的手段。

3. 小试样品的理化性质和生物学活性的初步检测　在实验室得到一定量经过纯化的目标产品后，需要对理化性质和生物学活性进行鉴定。理化性质一般包括对其分子量、等电点、纯度等进行测定，分析其氨基酸组成、N 末端及 C 末端的氨基酸序列，以及对其抗原性进行免疫学检测。根据活性蛋白的生物学性质，其生物学活性的检测可以采用体外细胞培养测定法、体内动物模型测定法，生化酶促反应测定法、离体动物器官测定法等手段进行。

在这个阶段中，各实验室需要遵守国家关于安全方面的规定，至于实验室的管理和试验操作规范方面没有统一的规定，每个实验室都有其各自的规定。

二、中试研究

中试生产是药物从实验室过渡到工业生产必不可少的重要环节，是二者之间的桥梁。

中试生产的主要任务是小试的扩大，是工业生产的缩影，应在工厂或专门的中试车间进行。中试研究的目的是建立一条完全模拟生产实际状况的小型的生产和质量控制的流水线进行中试生产。本阶段一般包括以下研究内容：

1. 生产工艺的研究　对于重组蛋白类产品的生产来说，一般需要经过发酵、粗纯、纯化、分装、冻干等阶段。在进行中试研究时，应对生产工艺的一系列参数和条件进行研究。通过对发酵条件、破菌方法、复性方法、纯化方案、色谱条件、制剂配方等进行研究，确定一套既合理、又经济，能生产稳定可靠的产品，并能放大至生产规模的工艺条件。

2. 质控方法和体系的研究和建立　在进行发酵、纯化、复性等一系列研究过程中，质控方法和体系的建立是工艺研究的前提保证。生物制品的一个显著特点是针对生产全过程的质量控制，因为有些指标无法在最终产品检测，即除对最终产品（原液和成品）进行完善的质量检测外，还应对整个生产过程进行在线检测和评价（on-line control or on-site evaluation），对生产中使用的原材料和辅料进行检测，对各个工艺环节的参数进行严格的控制。生物技术产品质控体系的建立分为以下几个主要方面：①用于生产重组 DNA 来源蛋白质产品的细胞或细菌表达载体的分析；②用于生物技术产品及生物制品生产的细胞基质的来源和鉴定；③生物技术产品稳定性试验；④最终产品和中间品的检测方法和质控标准的建立。

三、临床前研究

临床前研究阶段主要是对新药进行评选及对其生物学安全性进行临床前的评价，包括药理学、动物药代动力学和毒理学研究。

（一）临床前药理学研究

临床前药理学研究的主要目的是确证一个产品是否有防治或诊断疾病的作用，即是否有效。包括主要药效研究、一般药理研究、复方药理研究和药理作用机制研究等方面的工作。在生物技术药物进行申报时一般只涉及前两项研究的资料。评选新药一般总是从评价它的主要药效作用开始，从它预期用于临床预防、诊断、治疗目的的药理作用开始，这就是主要药效研究的任务。通过研究，明确受试药物的作用强度和特点，与老药相比，有什么优点可取，并从而决定有没有进一步系统评价的必要。在主要评价药效的同时，如有可能，还应努力阐明药物作用的作用部位和作用机制。一般药理学研究是指研究新药主要药效作用以外广泛药理作用的研究，尤其是对心血管系统、呼吸系统和神经系统的影响，为临床提供更多的信息。

（二）临床前药代动力学研究

药物代谢动力学（Pharmacokinetics），简称药代动力学，就是研究药物在机体内的代谢变化规律。临床前药代动力学就是研究药物在动物体内的吸收、分布、代谢和排泄的规律。生物技术药物由于其产品的特殊性，它们的药代动力学特点不同于传统小分子。动物种属间差别也将对生物技术药物药代动力学特点产生差别，因此需要选择合适的动物种属，并应考虑与药理学和毒理学研究中选用有关的动物种属，以保持资料的可比性。

（三）毒理学研究

新药毒理学研究的目的是确保临床用药安全。而要达到这一目的，就要完成急性毒性、长期毒性、特殊毒性和其他有关毒性试验。

1. 急性毒性试验　急性毒性试验（或单次给药的毒性试验）是观察动物接受单次或

短时间内（一般指 24 小时内）2～3 次给药后，动物出现的急性毒性反应。包括定性及定量实验。所谓定性试验，就是观察给药后动物有哪些中毒表现，出现和消失的时间过程，可能涉及哪些组织和器官，可能的靶器官是什么，分析中毒死亡原因等。所谓定量试验，主要是以死亡为评价的终点，测定药物引起动物死亡的剂量，包括致死量、近似致死量和半数致死量。

2. 长期毒性试验 所谓长期毒性试验（或重复给药的毒性试验）就是反复多次给药于动物，观察药物对动物的毒性反应，一般是指连续给药 14 天以上，具体给药期限的长短原则上应根据临床拟用的疗程而定。当完成了一个新药的主要药效学和急性毒性试验，并确认有进一步研究的价值后，才考虑进行长期毒性试验。其主要目的有二：一是观察在长期给药的情况下，实验动物出现的毒性反应、反应的剂量-效益关系、涉及哪些组织、主要的靶器官、损害程度及其可逆性等；二是观察在长期给药的情况下，实验动物能耐受的剂量范围及完全无毒的安全剂量，为判断候选的新药是否能进行临床试用、临床试用中应着重观测哪些生理生化指标以及选择临床初始剂量等提供动物试验的参考材料。

3. 特殊毒性试验 包括致突变试验、致癌试验、生殖毒性试验、免疫毒理学试验等。

4. 其他毒性试验 对一些特殊的药品，还包括其他的毒性试验，如皮肤给药毒性试验、腔道用药毒性试验、药物依赖性试验、抗生育药毒理研究和细胞毒抗肿瘤药毒理研究等。

四、临床研究

药物的临床研究包括临床试验和生物等效性试验。药物临床研究必须经国家食品药品监督管理局（state food and drug administration，SFDA）批准后实施，必须执行《药物临床试验质量管理规范》。临床试验分为Ⅰ、Ⅱ、Ⅲ、Ⅳ期。申请新药注册应当进行Ⅰ、Ⅱ、Ⅲ期临床试验，有些情况下可仅进行Ⅱ期和Ⅲ期，或者Ⅲ期临床试验。

临床试验各期的研究内容如下：

Ⅰ期临床试验：初步的临床药理学及人体安全性评价试验。观察人体对于新药的耐受程度和药代动力学，为制订给药方案提供依据。

Ⅱ期临床试验：治疗作用初步评价阶段。其目的是初步评价药物对目标适应证患者的治疗作用和安全性，也包括为Ⅲ期临床试验研究设计和给药剂量方案的确定提供依据。此阶段的研究设计可以根据具体的研究目的，采用多种形式，包括随机盲法对照临床试验。

Ⅲ期临床试验：治疗作用确证阶段。其目的是进一步验证药物对目标适应证患者的治疗作用和安全性，评价利益与风险关系，最终为药物注册申请获得批准提供充分的依据。试验一般应为具有足够样本量的随机盲法对照试验。

Ⅳ期临床试验：新药上市后应用研究阶段。其目的是考察在广泛使用条件下的药物的疗效和不良反应，评价在普通或者特殊人群中使用的利益与风险关系以及改进给药剂量等。

生物等效性试验，是指用生物利用度研究的方法，以药代动力学参数为指标，比较同一种药物的相同或者不同剂型的制剂，在相同的试验条件下，其活性成分吸收程度和速度有无统计学差异的人体试验。

药物临床试验的受试例数应当符合临床试验的目的和相关统计学的要求，并且不得少于《药品管理注册办法》附件规定的最低临床试验病例数。罕见病、特殊病种等情况，要求减少临床试验病例数或者免做临床试验的，应当在申请临床试验时提出，并经国家食品药品监督管理局审查批准。

临床试验用药物应当在符合《药品生产质量管理规范》的车间制备。制备过程应当严格执行《药品生产质量管理规范》的要求。申请人对临床试验用药物的质量负责。

五、新药临床研究与生产的申报和审批

（一）新药临床研究的申报与审批

按照《药品管理法》的规定，研制新药，必须按照国务院药品监督管理部门的规定如实报送研制方法、质量指标、药理及毒理试验结果等有关资料和样品，经国务院药品监督管理部门批准后，方可进行临床试验。该阶段的申报与审批的程序一般如下：

（1）申请人完成临床前研究后，应当填写《药品注册申请表》，向所在地省、自治区、直辖市药品监督管理部门如实报送有关资料。

（2）省、自治区、直辖市药品监督管理部门对申报资料进行形式审查，符合要求的，出具药品注册申请受理通知书；不符合要求的，出具药品注册申请不予受理通知书，并说明理由。

（3）省、自治区、直辖市药品监督管理部门自受理申请之日起5日内组织对药物研制情况及原始资料进行现场核查，对申报资料进行初步审查，提出审查意见。申请注册的药品属于生物制品的，还需抽取3个生产批号的检验用样品，并向药品检验所发出注册检验通知。

（4）省、自治区、直辖市药品监督管理部门在规定的时限内将审查意见、核查报告以及申报资料送交国家食品药品监督管理局药品审评中心，并通知申请人。

（5）接到注册检验通知的药品检验所按申请人申报的药品标准对样品进行检验，对申报的药品标准进行复核，并在规定的时间内将药品注册检验报告送交国家食品药品监督管理局药品审评中心，并抄送申请人。

（6）国家食品药品监督管理局药品审评中心收到申报资料后，在规定的时间内组织药学、医学及其他技术人员对申报资料进行技术审评，必要时可以要求申请人补充资料，并说明理由。完成技术审评后，提出技术审评意见，连同有关资料报送国家食品药品监督管理局。

（7）国家食品药品监督管理局依据技术审评意见作出审批决定。符合规定的，发给《药物临床试验批件》；不符合规定的，发给《审批意见通知件》，并说明理由。

（二）新药生产的申报与审批

新药在完成药物临床研究后，需要向国家食品药品监督管理局进行申报，经审批获得新药证书和药品批准文号后，才能进行生产和上市销售。该审批过程一般如下：

（1）申请人完成药物临床试验后，填写《药品注册申请表》，向所在地省、自治区、直辖市药品监督管理部门报送申请生产的申报资料，并同时向中国药品生物制品检定所报送制备标准品的原材料及有关标准物质的研究资料。

（2）省、自治区、直辖市药品监督管理部门应当对申报资料进行形式审查，符合要求的，出具药品注册申请受理通知书；不符合要求的，出具药品注册申请不予受理通知书，并说明理由。

（3）省、自治区、直辖市药品监督管理部门自受理申请之日起5日内组织对临床试验情况及有关原始资料进行现场核查，对申报资料进行初步审查，提出审查意见。除生物制品外的其他药品，还需抽取3批样品，向药品检验所发出标准复核的通知。省、自治区、直辖市药品监督管理部门在规定的时限内将审查意见、核查报告及申报资料送交国家食品药品监督管理局药品审评中心，并通知申请人。

（4）药品检验所应对申报的药品标准进行复核，并在规定的时间内将复核意见送交国家食品药品监督管理局药品审评中心，同时抄送通知其复核的省、自治区、直辖市药品监督管理部门和申请人。

（5）国家食品药品监督管理局药品审评中心收到申报资料后，应当在规定的时间内组织药学、医学及其他技术人员对申报资料进行审评，必要时可以要求申请人补充资料，并说明理由。

（6）经审评符合规定的，国家食品药品监督管理局药品审评中心通知申请人申请生产现场检查，并告知国家食品药品监督管理局药品认证管理中心；经审评不符合规定的，国家食品药品监督管理局药品审评中心将审评意见和有关资料报送国家食品药品监督管理局，国家食品药品监督管理局依据技术审评意见，作出不予批准的决定，发给《审批意见通知件》，并说明理由。

（7）申请人应当自收到生产现场检查通知之日起6个月内向国家食品药品监督管理局药品认证管理中心提出现场检查的申请。

（8）国家食品药品监督管理局药品认证管理中心在收到生产现场检查的申请后，应当在30日内组织对样品批量生产过程等进行现场检查，确认核定的生产工艺的可行性，同时抽取1批样品（生物制品抽取3批样品），送进行该药品标准复核的药品检验所检验，并在完成现场检查后10日内将生产现场检查报告送交国家食品药品监督管理局药品审评中心。

（9）样品应当在取得《药品生产质量管理规范》认证证书的车间生产；新开办药品生产企业、药品生产企业新建药品生产车间或者新增生产剂型的，其样品生产过程应当符合《药品生产质量管理规范》的要求。

（10）药品检验所应当依据核定的药品标准对抽取的样品进行检验，并在规定的时间内将药品注册检验报告送交国家食品药品监督管理局药品审评中心，同时抄送相关省、自治区、直辖市药品监督管理部门和申请人。

（11）国家食品药品监督管理局药品审评中心依据技术审评意见、样品生产现场检查报告和样品检验结果，形成综合意见，连同有关资料报送国家食品药品监督管理局。国家食品药品监督管理局依据综合意见，作出审批决定。符合规定的，发给新药证书，申请人已持有《药品生产许可证》并具备生产条件的，同时发给药品批准文号；不符合规定的，发给《审批意见通知件》，并说明理由。改变剂型但不改变给药途径，以及增加新适应证的注册申请获得批准后不发给新药证书；靶向制剂、缓释、控释制剂等特殊剂型除外。

第二节　医药生物技术产品研究开发的相关法规

医药生物技术产品研究开发的目的是为了提供用生物技术进行疾病的治疗、诊断和预防的新手段，为了确保所开发药品的质量，保证其安全、有效、可控，我国已建立了一整套的管理制度，下面介绍医药生物技术产品研究开发相关法规的主要内容，并重点介绍在临床前研究、临床研究以及生产阶段分别应遵守的质量管理规范。

一、医药生物技术产品研究开发相关法规的主要内容

为加强药品的管理，保证药品质量，增进疗效，保障人民用药安全有效，维护人民身体健康，国家制定了一系列法律、法规以及办法，其中与医药生物技术产品研究开发有关的法规和要求主要有这样三个部分。

1. 国家法律　包括《中华人民共和国药品管理法》和《中华人民共和国药品管理法实施条例》。《中华人民共和国药品管理法》于1984年9月20日第六届全国人民代表大会常务委员会第七次会议通过，2001年2月28日第九届全国人民代表大会常务委员会第二十次会议修订，自2001年12月1日起施行。《中华人民共和国药品管理法实施条例》于2002年8月4日由国务院公布，自2002年9月15日起施行。《中华人民共和国药品管理法》第二十九、三十条、三十一条以及《中华人民共和国药品管理法实施条例》第二十八条、二十九、三十条规定了新药的临床试验和生产都必须经国务院药品监督管理部门批准，药物非临床安全性评价研究机构必须执行《药物非临床研究质量管理规范》，药物临床试验机构必须执行《药物临床试验质量管理规范》。

2. 部门规章　从1985年至2000年，我国主管药品监督管理的部门先后制定一系列新药管理的法规和制度。主要有《新药审批办法》、《新生物制品审批办法》、《新药保护和技术转让的规定》、《仿制药品审批办法》、《进口药品管理办法》和《药品注册管理办法》等，这些是《中华人民共和国药品管理法》的重要配套规章。其中新的《药品注册管理办法》于2007年6月18日经国家食品药品监督管理局局务会审议通过，自2007年10月1日起施行，是目前对药品注册进行规范化管理的主要依据。该“办法”是在总结我国多年药品注册管理经验的基础上，借鉴了国外先进的注册管理方法制定的。它适应了新的《药品管理法》及其《实施条例》对药品注册管理的要求，适应了中国加入WTO以后的新的医药经济形势。新的《药品注册管理办法》共15章177条，包括以下内容：

第一章　总则
第二章　基本要求
第三章　药物的临床试验
第四章　新药的申报与审批
　第一节　新药临床试验
　第二节　新药生产
　第三节　新药监测期
第五章　仿制药的申报与审批
第六章　进口药品的申报与审批
　第一节　进口药品的注册
　第二节　进口药品分包装的注册
第七章　非处方药的申报
第八章　补充申请的申报与审批
第九章　药品再注册
第十章　药品注册检验
第十一章　药品注册标准和说明书
　第一节　药品注册标准
　第二节　药品标准物质
　第三节　药品名称、说明书和标签
第十二章　时限
第十三章　复审
第十四章　法律责任
第十五章　附则

另外，对药品研制、临床和生产全过程，国家食品药品监督管理局已经出台规范性的要求，它们是《药物非临床研究质量管理规范》(Good Laboratory Practice，GLP)、《药品

临床试验管理规范》（Good clinical practice，GLP）和《药品生产质量管理规范》(Good manufacturing practice，GMP)。这在前面已有介绍。

3. 技术指导性文件　为帮助和指导新药研制单位用科学、规范的方法开展新药研究申报工作，参照世界发达国家的标准，结合我国的实际，国家药品监督管理部门组织专家对新药技术要求进行了详细的规定。它们虽不是法规，但根据这些规定去做得到可靠的研究结果，同时也是新药审批时的依据，因此对于新药的研究十分重要。

近几年已公布执行的有：①新药（西药）临床前药学、药理学和毒理学评价的指导原则；②新药（西药）临床研究指导原则；③中药新药研究指南，其中包括了中药新药临床前药学、药理学和毒理学评价指南。2003 年 3 月 20 日，国家食品药品监督管理局印发了《预防用以病毒为载体的活疫苗制剂的技术指导原则》、《预防用 DNA 疫苗临床前研究技术指导原则》、《艾滋病疫苗临床研究技术指导原则》、《人用重组 DNA 制品质量控制技术指导原则》、《人用单克隆抗体质量控制技术指导原则》、《人基因治疗研究和制剂质量控制技术指导原则》、《人体细胞治疗研究和制剂质量控制技术指导原则》、《变态反应原(变应原)制品质量控制技术指导原则》和《细胞培养用牛血清生产和质量控制技术指导原则》等 9 个技术指导原则。2005 年 10 月 14 日，国家食品药品监督管理局印发了《预防用疫苗临床前研究技术指导原则》、《生物制品生产工艺过程变更管理技术指导原则》、《联合疫苗临床前和临床研究技术指导原则》、《多肽疫苗生产及质控技术指导原则》、《结合疫苗质量控制和临床研究技术指导原则》和《预防用疫苗临床试验不良反应分级标准指导原则》等 6 个技术指导原则。

二、药物非临床研究质量管理规范

《药物非临床研究质量管理规范》对应的英文是“Good Laboratory Practice for Non-clinical Laboratory Studies”或“Non-clinical Good Laboratory Practice”，简称 GLP。GLP 是为申请药品注册而进行的非临床研究必须遵守的规定。非临床研究，系指为评价药物安全性，在实验室条件下，用实验系统进行的各种毒性试验，包括单次给药的毒性试验、反复给药的毒性试验、生殖毒性试验、遗传毒性试验、致癌试验、局部毒性试验、免疫原性试验、依赖性试验、毒代动力学试验及与评价药物安全性有关的其他试验。1979 年美国颁布 GLP 后，许多国家和国际组织制定和颁布了各自的 GLP，其内容大同小异。GLP 逐渐成为国与国之间相互承认新药的基础。在我国，国家食品药品监督管理局于 2003 年 8 月 6 日发布《药物非临床研究质量管理规范》，自 2003 年 9 月 1 日起实施。其主要内容有：

1. 组织机构和工作人员　非临床安全性评价研究机构应建立完善的组织管理体系，配备机构负责人、质量保证部门负责人和相应的工作人员。明确规定了非临床安全性评价研究机构人员的要求，明确规定了机构负责人、质量保证部门负责人和专题负责人各自的职责。另外明确规定了非临床安全性评价研究机构应设立独立的质量保证部门。

2. 实验设施

(1) 根据所从事的非临床研究的需要，建立相应的实验设施。各种实验设施应保持清洁卫生，运转正常；各类设施布局应合理，防止交叉污染；环境条件及其调控应符合不同设施的要求。

(2) 具备设计合理、配置适当的动物饲养设施，并能根据需要调控温度、湿度、空气洁净度、通风和照明等环境条件。实验动物设施条件应与所使用的实验动物级别相符。

(3) 具备饲料、垫料、笼具及其他动物用品的存放设施。各类设施的配置应合理，防止与实验系统相互污染。易腐败变质的动物用品应有适当的保管措施。

（4）具有供试品和对照品的处置设施。

（5）根据工作需要设立相应的实验室；使用有生物危害性的动物、微生物、放射性等材料应设立专门实验室，并应符合国家有关管理规定。

（6）具备保管实验方案、各类标本、原始记录、总结报告及有关文件档案的设施。

（7）根据工作需要配备相应的环境调控设施。

3. 仪器设备和实验材料

（1）根据研究工作的需要配备相应的仪器设备，放置地点合理，并有专人负责保管，定期进行检查、清洁保养、测试和校正，确保仪器设备的性能稳定可靠。

（2）实验室内应备有相应仪器设备保养、校正及使用方法的标准操作规程。对仪器设备的使用、检查、测试、校正及故障修理，应详细记录日期、有关情况及操作人员的姓名等。

（2）供试品和对照品的管理应符合要求。

（3）对实验室的试剂和溶液、动物的饲料和饮水、动物饲养室内使用的清洁剂、消毒剂及杀虫剂等的管理应符合要求。

4. 标准操作规程　应制订与实验工作相适应的标准操作规程，并规定了需要制订标准操作规程的范围。标准操作规程的制订、批准、修改、分发、存放、销毁等应符合规定。研究过程中任何偏离标准操作规程的操作，都应经专题负责人批准，并加以记录。标准操作规程的改动，应经质量保证部门负责人确认，机构负责人书面批准。

5. 研究工作的实施

（1）专题负责人应制订实验方案，经质量保证部门审查，机构负责人批准后方可执行，批准日期作为实验的起始日期。接受委托的研究，实验方案应经委托单位认可。

（2）专题负责人全面负责研究专题的运行管理。参加实验的工作人员，应严格执行实验方案和相应的标准操作规程，发现异常现象时应及时向专题负责人报告。

（3）研究工作结束后，专题负责人应及时写出总结报告，签名或盖章后交质量保证部门负责人审查和签署意见，机构负责人批准。

6. 资料档案　研究工作结束后，专题负责人应将实验方案、标本、原始资料、文字记录和总结报告的原件、与实验有关的各种书面文件、质量保证部门的检查报告等按标准操作规程的要求整理交资料档案室，并按标准操作规程的要求编号归档。

7. 监督检查

（1）国家食品药品监督管理局负责组织实施对非临床安全性评价研究机构的检查。

（2）凡为在中华人民共和国申请药品注册而进行的非临床研究，都应接受药品监督管理部门的监督检查。

三、药物临床试验质量管理规范

《药物临床试验质量管理规范》对应的英文是 Good Clinical Practice（GCP）。1964 年世界医生总会发表“赫尔辛基宣言”，对以人体作为生物医学研究的医务人员，提出了伦理和科学标准方面的要求。宣言引起广泛注意，部分研究开发新药多的国家对新药临床研究管理制定了指南或规范。中国的《药物临床试验质量管理规范》是 2003 年 8 月 6 日由国家食品药品监督管理局发布，自 2003 年 9 月 1 日起施行，共 13 章 70 条，其主要内容有：

1. 临床试验前的准备与必要条件　进行药物临床试验必须有充分的科学依据。在进行人体试验前，必须周密考虑该试验的目的及要解决的问题，应权衡对受试者和公众健康

预期的受益及风险，预期的受益应超过可能出现的损害。选择临床试验方法必须符合科学和伦理要求。

2. 受试者的权益保障 药物临床试验的过程中，必须对受试者的个人权益给予充分的保障，并确保试验的科学性和可靠性。受试者的权益、安全和健康必须高于对科学和社会利益的考虑。伦理委员会与知情同意书是保障受试者权益的主要措施。

3. 试验方案 临床试验开始前应制订试验方案，该方案应由研究者与申办者共同商定并签字，报伦理委员会审批后实施。

4. 研究者的职责 负责临床试验的研究者应具备下列条件：

在医疗机构中具有相应专业技术职务任职和行医资格；

具有试验方案中所要求的专业知识和经验。

对临床试验方法具有丰富经验或者能得到本单位有经验的研究者在学术上的指导；

熟悉申办者所提供的与临床试验有关的资料与文献；

有权支配参与该项试验的人员和使用该项试验所需的设备。

5. 申办者的职责 申办者负责发起、申请、组织、监察和稽查一项临床试验，并提供试验经费。申办者按国家法律、法规等有关规定，向国家食品药品监督管理局递交临床试验的申请，也可委托合同研究组织执行临床试验中的某些工作和任务。

6. 监查员的职责 监查员是申办者与研究者之间的主要联系人。其人数及访视的次数取决于临床试验的复杂程度和参与试验的医疗机构的数目。监查员应有适当的医学、药学或相关专业学历，并经过必要的训练，熟悉药品管理有关法规，熟悉有关试验药物的临床前和临床方面的信息以及临床试验方案及其相关的文件。监查员应遵循标准操作规程，督促临床试验的进行，以保证临床试验按方案执行。

7. 记录与报告 病历作为临床试验的原始文件，应完整保存。病例报告表中的数据来自原始文件并与原始文件一致，试验中的任何观察、检查结果均应及时、准确、完整、规范、真实地记录于病历和正确地填写至病例报告表中，不得随意更改，确因填写错误，作任何更正时应保持原记录清晰可辨，由更正者签署姓名和时间。临床试验中各种实验室数据均应记录或将原始报告复印件粘贴在病例报告表上，在正常范围内的数据也应具体记录。对显著偏离或在临床可接受范围以外的数据须加以核实。检测项目必须注明所采用的计量单位。为保护受试者隐私，病例报告表上不应出现受试者的姓名。研究者应按受试者的代码确认其身份并记录。

8. 数据管理与统计分析 临床试验中受试者分配必须按试验设计确定的随机分配方案进行，每名受试者的处理分组编码应作为盲底由申办者和研究者分别保存。临床试验资料的统计分析过程及其结果的表达必须采用规范的统计学方法。临床试验各阶段均需有生物统计学专业人员参与。临床试验方案中需有统计分析计划，并在正式统计分析前加以确认和细化。若需作中期分析，应说明理由及操作规程。对治疗作用的评价应将可信区间与假设检验的结果一并考虑。所选用统计分析数据集需加以说明。对于遗漏、未用或多余的资料须加以说明，临床试验的统计报告必须与临床试验总结报告相符。

9. 试验用药品的管理 临床试验用药品不得销售。试验用药品的使用记录应包括数量、装运、递送、接受、分配、应用后剩余药物的回收与销毁等方面的信息。

10. 质量保证 申办者及研究者均应履行各自职责，并严格遵循临床试验方案，采用标准操作规程，以保证临床试验的质量控制和质量保证系统的实施。

11. 多中心试验 多中心试验是由多位研究者按同一试验方案在不同地点和单位同时进行的临床试验。各中心同期开始与结束试验。多中心试验由一位主要研究者总负责，并作为临床试验各中心间的协调研究者。

四、药品生产质量管理规范

《药品生产质量管理规范》是在药品生产全过程中，保证生产出优质药品的管理制度。美国最先通过立法制定了药品生产管理规范，其原名为“Good Practice in the Manufacturing and Quality Control of Drugs”，简称“Good Manufacturing Practice”。以后人们均简称此制度为“GMP”。1982年，中国医药工业公司和中国药材公司分别制定了《药品生产管理规范(试行)》、《中成药生产质量管理办法》，这是我国制药工业组织制定的GMP，也是我国最早的GMP。1984年颁布的《中华人民共和国药品管理法》第9条规定：“药品生产企业必须按照国务院卫生行政主管部门制定的《药品生产质量管理规范》的要求，制定和执行保证药品质量的规章制度和卫生要求”。根据这一法律规定，卫生部于1988年颁布了我国法定的《药品生产质量管理规范》。1992年颁布修订版，1998年国家食品药品监督管理局再次颁布修订版。

GMP是药品生产过程质量管理实践中总结、抽象、升华出来的规范化的条款，它的目的是为了指导药品生产企业克服不良生产导致劣质药品产生，保证优质生产合格药品。它的覆盖面是所有药品、所有药品生产企业。

GMP的内容很广泛，人们从不同角度来概括其内容。从专业性管理的角度，可以把GMP分为两大方面。一是对原材料、中间品、产品的系统质量控制，主要办法是对这些物质的量进行检验，并随之产生了一系列工作质量管理。另一方面是对影响药品质量的，生产过程中易产生的人为差错和污染异物引入，进行系统严格管理，以保证生产合格药品。前者称为质量管理，后者称为质量保证。从硬件和软件系统的角度，可以将GMP分为硬件系统和软件系统。硬件系统主要包括人员、厂房、设施、设备等的目标要求，这部分涉及必需的人财物的投入，以及标准化管理。软件系统主要包括组织机构、组织工作、生产工艺、记录、制度、方法、文件化程序、培训等，可以概括为以智力投入为主的投入产出。

我国的GMP（1998年修订）有14章89条：第一章总则；第二章机构与人员；第三章厂房与设施；第四章设备；第五章物料；第六章卫生；第七章验证；第八章文件；第九章生产管理；第十章质量管理；第十一章产品销售与收回；第十二章投诉与不良反应报告；第十三章自检；第十四章附则。

第一章总则：包括2条。说明了制定GMP的根据，同时规定GMP是药品生产和质量管理的基本准则，规定GMP适用于制剂的全过程和原料药生产中影响成品质量的关键工序。

第二章：包括5条。规定药品生产企业应建立生产管理和质量管理机构；企业主管生产和质量管理的负责人应具有医药或相关专业大专以上学历，具有药品生产和质量管理的经验，对规范的实施和产品质量负责。药品生产管理和质量管理部门的负责人应具有医药或相关专业大专以上学历，有药品生产管理和质量管理的实践经验，有能力对药品生产和质量管理中的实际问题作出正确判断和处理。

第三章：包括23条。规定企业必须有整洁的环境，厂房防止昆虫和其他动物进入，便于清洁。洁净室（区）内表面应平整光滑、无裂缝、接口严密、无颗粒脱落、地面与墙交界处应呈弧形。洁净室照度宜为300Lux，温度18℃～26℃，相对湿度控制在45%～65%。对青霉素类、β-内酰胺类、避孕药、激素类、抗肿瘤类的净化和设备做了相应的规定，防止交叉污染。

第四章：包括7条。对设备的设计、选型安装、维修、保养和验证作出规定，并有明

显的状态标识。规定注射用水的储存可采用 80℃以上保温、65℃以上循环保温或 4℃以下存放。规定生产和检验用的仪器、仪表、量具、衡器精密度应符合生产和检验要求，定期校验，有明显的合格标志。

第五章：包括 10 条。规定必须有物料购入、储存、发放、使用等管理制度；所用物料必须符合标准；物料按规定条件储存，有明显的检验状态标志；物料按使用期限储存，无使用期限的，储存期不超过 2 年；药品的标签和使用说明书由专人保管、领用。

第六章：包括 9 条。规定要制订各项卫生制度；各岗位要制订清洁规程；无菌工作服必须包盖全部头发、胡须及脚部，并能阻留人体脱落物；进入洁净室人员不得化妆和佩带饰物，不得裸手直接接触药品。

第七章：包括 4 条。规定药品生产验证包括厂房、设施及设备安装确认、运行确认和产品验证；当影响产品质量的主要因素发生改变时，以及生产一定生产周期后应进行再验证。

第八章：包括 5 条。规定了生产管理和质量管理的各项制度和记录；规定生产管理文件有生产工艺规程、岗位操作法和标准操作规程；批生产记录；规定质量管理文件；制订文件的要求等。

第九章：包括 8 条。规定每批产品应按产量和数量的物料平衡进行检查；规定具有同一性质和质量，并在同一连续生产周期中生产出来的一定数量的药品为一批，每批药品均应编制生产批号；药品包装应有记录；每批药品每一生产阶段完成后必须由生产操作人员清场，填写清场记录。

第十章：包括 3 条。规定每批产品均应有销售记录，可进行追踪；退货或收回应有记录。

第十二章：包括 3 条。规定应建立不良反应监察报告制度。

第十三章：2 条。规定企业应定期组织自检，自检应有记录。

第十四章：3 条。对 11 条用语确定其含义。

第三节　医药生物技术及其产品的生物安全问题

广义的“生物安全”是指在一个特定的时空范围内，由于自然或人类活动引起的外来物种迁入，并由此对当地其他物种和生态系统造成改变和危害；人为造成环境的剧烈变化而对生物的多样性产生影响和威胁；在科学研究、开发、生产和应用中造成对人类健康、生存环境和社会生活有害的影响。凡此上述均属于“生物安全”的范畴。国内外由于生物安全问题造成对人类以及对生态环境危害的事例不胜枚举，这里讨论通过医药生物技术特别是基因工程技术所产生的转基因生物安全性问题。

一、医药生物技术及其产品生物安全问题的范围

医药生物技术及其产品存在潜在的危险性，无意中构建出危及人类安全的微生物的可能性仍然是不能排除的；由现代生物技术产生，而又难以完全纳入现行药品管理法规，可能对人类健康和生态环境产生潜在的直接或间接危害的因素有：重组活疫苗、DNA 疫苗、生物治疗包括基因治疗和体细胞治疗、反义核酸药物、以生产医药产品为目的的转基因动植物等。这迫使人们始终关注着重组 DNA 技术对人类及其环境有可能带来的安全性问题，重组修饰的生物是否对人类和其他生物有害？它们是否会在环境中极度繁殖造成危害？重组活疫苗、反义核酸药物、基因治疗药物等医药生物技术产品进入人体后是否具有危

害性？

关于医药生物技术及其产品的生物安全问题主要包括以下三个方面：

1. 实验室重组 DNA 操作的潜在危害　实验室重组 DNA 操作的潜在危害主要表现在两个方面：一是实验室病原体或重组病原体感染操作者所造成的实验室性感染；二是带有重组 DNA 的载体或受体的动植物、细菌及病毒逃逸出实验室造成社会性污染。

实验室性感染的途径很多，如操作者体表污染未能及时清除、在实验室进食将实验微生物带进消化道、实验操作失误导致创口感染等都是常见的感染方式，此外，实验材料形成的气溶胶颗粒进入实验人员的呼吸道也是造成实验室性感染的重要原因之一，这是一种不易察觉和防范的感染方式。实验室性感染一方面在于危害实验室工作人员的身体健康；另一方面若实验室病原体通过操作者的社会活动带至实验外扩散，有可能进一步危害社会。

2. 基因工程工业化生产的潜在危害　大规模基因工程工业化生产所涉及的安全性问题比实验室中进行重组 DNA 试验更为复杂。主要有：①感染危险，即由于接触活菌体或病毒而使人、动物及植物可能发生疾病；②生产过程中的死菌体或死细胞及其组分或代谢产物对人体及其他生物造成的毒性、致敏性及其他生物学效应；③产品的毒性、致敏性及其他生物学效应；④环境效应。在工业生产中使用遗传工程体，由于体积大、密度高和持续时间较长以及操作人员所受教育程度的参差不齐，使发生危害的可能性比小规模的实验室工作要大得多，并且，一旦出现问题，后果也会更加严重。

3. 生物技术产品的安全性问题　目前基因工程、疫苗和单克隆抗体已有大量上市，新的生物技术产品如包括基因治疗，DNA 药物，组织工程产品等有如雨后春笋不断涌现。这些近十年发展起来的与传统医药不同的治疗方法，由于其临床经验少，科学发展的阶段性所限，对于外源基因进入人体后长远的影响目前尚不能完全了解，国内外研究参考文献也很有限，对这些新类型的药物究竟对人体及环境有何不良影响作用，我们需要非常慎重地对待。

二、医药生物技术及其产品生物安全问题的预防和控制

20 多年的实践证明，重组 DNA 操作的潜在危害是可以采用适当措施加以防止的。第一，要求从事重组 DNA 操作的实验人员具备良好的从事微生物操作的能力以及关于安全防护的基本知识，并正确认识实验生物的危害等级以及有关的重组 DNA 工作的类型，从而采用不同的操作技术和封闭措施；第二，尽可能选择具有生物控制功能的宿主-载体系统；第三，加强立法管理重组 DNA 试验，是保障生物安全的重要环节。

为了预防医药生物技术产品的潜在危险，目前各国医药生物技术的安全管理均纳入现行药品管理体系。近年来世界卫生组织（world health organization，WHO）、国际药品注册协调组织（the international conference on harmonization of technical requirements for registration of pharmaceutical for human use，ICH）、美国食品和药品管理局（food and drug administration，FDA）颁布一系列有关生物技术药物的生产和检定规程、技术指南及质量控制要点。根据新药安全、有效、可控的原则要求，以个案处理的方式（case by case），具体评价每一个产品或治疗方案的安全性，其项目除常规新药必须进行的急性毒性试验、亚急性毒性试验、长期毒性试验、特殊毒性试验外，还必须考察由于采用现代生物技术而可能产生的特殊安全性问题，如基因治疗、DNA 疫苗、反义核酸药物的不慎应用，或因纯化不彻底在基因工程产品中残留有较高水平的宿主 DNA 或病原体如支原体和病毒等，都有可能导致人体基因的异常。对于重组 DNA 试验的安全控制措施、重组 DNA

药物和单克隆抗体质量控制要点、DNA 疫苗质量保障指南以及体细胞治疗和基因治疗考虑要点等都具有相应的安全管理措施，而转基因动植物生产医药产品，还必须符合转基因动植物生产的安全管理规定。

为了加强对包括生物技术产品在内的医药产品的安全管理，国务院于 1998 年专门成立了国家食品药品监督管理局，下设中国药品生物制品检定所、药品审评中心、药品认证管理中心、国家药典委员会、国家中药品种保护审评委员会、药品评价中心等，负责对各种医药产品的研究、生产、流通和使用的安全和质量监督管理。于 1998 年重新修订了我国医药行业的《药品生产质量管理规范》(GMP)，保证包括医药生物技术产品在内的药品生产全过程中对人员和内外环境的安全和产品的质量。重新修订了新的医药生物技术产品审批办法，加强对医药生物技术产品在研制过程中的资料审查，制订了严格的临床前安全和毒性试验检查方案和方法，成立了各种疫苗、多肽药物、基因治疗和体细胞治疗等专门专家评审委员会。制订了若干个对各类制品的安全和质量具有指导性的技术指南。这些文件在保障医药生物技术产品在研究、开发、生产和使用过程中对内外环境和人体的安全、效果等方面起到了极为重要的作用。

(王军志　李永红)

参考书目

1. 贾弘禔. 生物化学. 北京：人民卫生出版社，2005
2. 查锡良. 生物化学. 北京：人民卫生出版社，2008
3. 冯作化. 医学分子生物学. 北京：人民卫生出版社，2005
4. 查锡良. 医学分子生物学. 北京：人民卫生出版社，2004
5. J. D. 沃森. 基因的分子生物学. 北京：科学出版社，2005
6. David LN, Michael MC. Lehninger Principles of Biochemistry. 4th ed. New York: Worth Publishers, 2005
7. Benjiamin Lewin, Gene IX. 9th ed. Sudbury: Jones & Bartlett Publishers, 2007
8. David Clark. Molecular Biology. 北京：科学出版社，2007
9. Tom Strachan, Andrew P Read. Human Molecular Genetics. New York: Carland Publishing, 2004
10. Robert F. Weaver. Molecular Biology. 2th ed. New York: McGraw-Hill Higher Education, 2002
11. Jorde LB, Carey JC, White RL. Medical Genetics. 2th ed. St. Louis: Mosby, 2000
12. 黄汉菊. 医学微生物学. 北京：高等教育出版社，2002
13. Brooks GF, Butel JS and Morse SA. Jawetz, Melnick, & Adelberg's Medical Microbiology. 23th ed. New York: McGraw Hill, 2004
14. 李玉林. 病理学. 第 6 版. 北京：人民卫生出版社，2005
15. 何维. 医学免疫学. 北京：人民卫生出版社，2006
16. 金惠铭. 病理生理学. 第 6 版. 北京：人民卫生出版社，2005
17. Chien K R. 分子心脏病学. 北京：人民卫生出版社，2002
18. 李小鹰. 心血管疾病分子生物学. 北京：人民军医出版社，2000
19. David G. Gardner, Dolores Shoback. Greenspan's Basic & Clinical Endocrinology. 8th ed. New York: McGraw-Hill, 2007
20. 高晓明. 医学免疫学基础. 北京：北京大学医学出版社，2005
21. 童坦君，张宗玉. 医学老年学-衰老与长寿. 北京：人民卫生出版社，1995
22. 刘焯霖，梁秀龄，张成. 神经遗传病学. 北京：人民卫生出版社，2002
23. Charles R. Scriver, Arthur L. Beaudet, William S. Sly, et al. The Metabolic & Molecular Basis of Inherited Disease. 8th ed. New York: McGraw-Hill, 2007
24. Cox TM, Scinclair J. Molecular Biology in Medicine. Oxford: Blackwell Science Limited, 2000
25. Trent R. J. Molecular Medicine. 3th ed. Burlington: Elsevier Academic Press, 2005
26. 陆国辉，徐湘民. 临床遗传咨询. 北京：北京大学医学出版社，2007
27. 马大龙. 生物技术药物. 北京：科学出版社，2001
28. 焦炳华. 现代生物工程. 北京：科学出版社，2007
29. The 2000—2005 World Outlook for Biotechnology, Icon Group Ltd., 2002

30. 顾健人，曹雪涛. 基因治疗. 北京：科学出版社，2002
31. 查锡良. 医学分子生物学. 北京：人民卫生出版社，2004
32. 徐文方. 新药设计与开发. 北京：科学出版社，2001
33. 王军志. 生物技术药物研究开发和质量控制. 第 2 版. 北京：科学出版社，2007

中文索引

3′端非翻译区 133
α-共核蛋白 223
σ亚基 57

A

Aβ淀粉样前体蛋白 221
阿尔茨海默病 220
癌基因 140
氨基酰tRNA合成酶 66
氨基酰位或A位 69

B

B细胞受体 195
*BRCA*基因 55
靶点 299
半不连续复制 30
胞质小RNA 19
比较基因组学 88
闭合转录复合体 58
变性高效液相色谱 263
变异 128
表达序列标签 90
表达载体 233
表观遗传 82
表皮生长因子 118
表型克隆 249
表型转化 178
病毒癌基因 141
病毒吞饮 155
病理性衰老 216
病原体相关的分子模式 162
病原性 150
补体受体3 164
不对称转录 57

C

CDK抑制蛋白 117
cDNA微阵列 249
cDNA文库 231
cDNA芯片 240
操纵序列 77
操纵元件 25
操纵元件或称操纵序列 14
操纵子 25,76
差异基因表达 75
产前诊断 267
长QT综合征 183
长分散重复片段 23
长末端重复序列 42,288
长寿基因 217
超抗原 157
超螺旋 9
沉默复合物(RISC 244
沉默子 16
成体干细胞 125
程序衰老学说 217
穿梭载体 236
促有丝分裂因子 117

D

DNA甲基化 82
DNA结合结构域 83
DNA聚合酶 29,228
DNA克隆 226
DNA连接酶 29,233
DNA酶 19
DNA芯片 262
DNA序列测定 236
DNA依赖的蛋白激酶 51
DNA指导的RNA聚合酶 57
DNA指纹 260
DNA重组 226
大片段突变 129
代表性差异分析 249
代谢组学 88,304
单倍基因型 260
单倍型 131,202,260
单纯疱疹病毒 288

单核苷酸多态性 90,128
单拷贝或低度重复序列 22
单链 DNA 结合蛋白 29
单顺反子 72
蛋白激酶 100,212
蛋白激酶偶联受体 100
蛋白酪氨酸激酶 100,197
蛋白酶体 116
蛋白质组 93
蛋白质组学 88
等位基因特异性寡核苷酸 262
低密度脂蛋白 180
低氧诱导因子-1 213
第二信使 100
点突变 129
电子克隆 253
定位克隆 251
动脉粥样硬化 177
毒力 149
端粒 12,37
端粒酶 38
短串联重复突变 129
短串联重复序列 90
短分散重复片段 23
断裂基因 12
多泛素化 116
多基因家族 24
多克隆位点 233
多能性 124
多顺反子 mRNA 25,72

F

翻译 12,56
反式作用因子 13,75
反义 RNA 19,278,285
反义技术 277
反转录 228
反转录-PCR 231
反转录病毒 42,288
反转录酶 23
泛素 116
非翻译区 17
非同源末端连接重组 51
分解代谢物基因活化蛋白 78
分子伴侣 71
分子克隆 226
辅助病毒 289
复制叉 30
复制起点 29
复制许可因子 39
复制子 30,233

G

G 蛋白 101
G 蛋白偶联受体 100
G 蛋白循环 106
钙调蛋白 108
钙离子依赖的细胞黏附素家族 164
感染性疾病 149
冈崎片段 31
高度重复序列 22
高迁移率族蛋白 B1 162
高通量筛选 301
高脂血症 180
功能蛋白质组学 95
功能获得性突变 131
功能基因组学 88
功能克隆 248
功能失去性突变 131
寡核苷酸 277
管家基因 75
光复活酶 48
过氧化物酶体增殖因子激活受体-γ 194

H

合理药物设计 297
和血管紧张素Ⅱ 176
核酶 19,278
核酸 6
核酸分子杂交 227
核酸内切酶 19
核酸外切酶 19
核酸疫苗 279
核糖核酸 6
核糖体 18
核糖体 RNA 12

核糖体蛋白 18
核糖体循环 69
核小 RNA 19
核小核糖核蛋白颗粒 19
核小体 10
核心酶 57
横向分化或转分化 121
后基因组学 88
互补 DNA 23
化学基因组学 304
环氧化酶 168
活性氧 44
霍乱毒素 158

J

肌营养不良蛋白 253
基因表达 12,56
基因表达产物 240
基因表达的系列分析 249
基因表达调控 74
基因表达系列分析 92
基因操作 226
基因打靶 243
基因定位 251
基因定向突变 238
基因工程 226
基因工程亚单位疫苗 279
基因工程疫苗 279
基因鉴定集成法 249
基因矫正 285
基因克隆 233
基因敲除 243
基因敲入 243
基因添加 285
基因芯片 92
基因型 6,131
基因药物 277
基因疫苗 279
基因诊断 257
基因治疗 284
基因置换 285
基因转染 236
基因转移 285
基因组 20
基因组错配筛选 249
基因组构 12
基因组文库 231
基因组学 88
基于机制的药物设计 300
急性期蛋白 208
疾病相关基因 246
疾病易感性 128
继发性免疫缺陷病 195
假肥大型肌营养不良 252
间接诊断 258
间接体内疗法 286
剪接体 62
剪切和多聚 A 作用特异因子 62
碱基对 9
碱性亮氨酸拉链 84
碱性螺旋-环-螺旋 83
降钙素基因相关肽 168
酵母人工染色体 90
酵母双杂交系统 255
阶段特异性 74
结构蛋白质组学 95
结构基因 12
解旋酶 29
聚丙烯酰胺凝胶电泳 236
聚合酶链式反应 228

K

Kozac 共有序列 68
开放阅读框 16,65
开放转录复合体 58
抗谷氨酸脱羧酶 204
拷贝数变异 128
拷贝数多态性 128
可诱导型表达 242
克隆载体 233
空间特异性 75
孔道蛋白 153
扩增片段长度多态性 260
扩增子 250

L

酪氨酸激酶 JAK 110

连锁不平衡 259
连锁分析 252,258
连锁图 89
链激酶 274
淋巴细胞功能相关抗原 1 164
磷脂酶 C 100
磷脂酰肌醇-3-激酶 110
流出泵 153

M

mRNA 差异显示 249
帽结合蛋白 17
密码子 17,65
免疫球蛋白 195
免疫球蛋白超家族 164
免疫缺陷综合征 159
免疫受体酪氨酸激活模体 197
模板链 57
模体 83

N

Northern 印迹 240
内毒素 150
内含子 12
内皮素-1 176
逆向药理学 298
尿激酶 274

P

PI3K 相关激酶 211
poly(A)结合蛋白 17
poly(A)聚合酶 62
P 物质 168
帕金森病 220
旁抑制效应 122
胚胎干细胞 123,242
配体门控受体型离子通道 100
平齐末端 234
葡激酶 274
普里布诺盒 14

Q

气体信号分子 177
启动子 14,25,82
起始密码子 66
起始识别复合物 36
千碱基对 9
前列腺素 168
强直性脊柱炎 204
趋化性迁移 165
趋化因子 165
全酶 57
全能性 123

R

RNA 编辑 64
RNA 干扰 244
RNA 酶 19
RNA 酶保护试验 237
RNA 依赖的 RNA 聚合酶 244
RNA 诱导 244
RNA 诱导的沉默复合物 86
RNA 组学 19
染色体 10
染色体异常 252
染色体原位杂交 251
染色质 10
热休克蛋白 71,133,209
热休克因子 210
热休克元件 210
人类白细胞抗原 195
人类基因组计划 88
人类免疫缺陷病毒 159
乳糖操纵子 77

S

SD 序列 69
SH2 结构域 101
SH3 结构域 101
SNP 的单倍型分析 260
Southern 印迹 239
三联体密码 17
三螺旋 DNA 9
色氨酸操纵子 78
杀菌性/通透性增强蛋白 276
神经生长因子 118

神经退行性疾病 220
生长因子 118
生理性衰老 216
生物技术 226
生物信息学 92
生殖系突变 128
时间特异性 74
实时 PCR 239
实时 RT-PCR 240
视网膜母细胞瘤 143
释放因子 70
噬菌体展示 255
受体 100
受体连结蛋白 155
树突状细胞 293
衰减子 80
衰老 216
衰老基因 217
双螺旋 8
双脱氧核苷三磷酸 236
顺式调节 13
顺式作用元件 13,75
瞬时转染 242
丝裂原激活的蛋白激酶 109
羧基末端结构域 58

T

TATA 盒 15
Toll 样受体 161
T 细胞受体 195
肽核酸 278
肽酰位或 P 位 69
探针 239
糖组学 88
特异转录因子 85
体内疗法 285
体细胞突变 128
体细胞杂交 251
调节基因 76
调控区 12
条件性基因敲除 244
通用转录因子 85
同源表达序列标签 253
同源重组 51
突变分析 258
退火 227
脱氧核糖核苷三磷酸 29
脱氧核糖核酸 6
拓扑异构酶 29

W

Western 印迹 240
外毒素 150
外显子 12
微 RNA 19,86
微卫星 90
卫星 DNA 22
稳定转染 242
稳定转染细胞 254
物理图 90

X

X 连锁高 IgM 血症 198
X 连锁无丙种球蛋白血症 198
系统生物学 96,305
系统性红斑狼疮 205
细胞癌基因 141
细胞分化 120
细胞分裂后期促进复合体 117
细胞分裂周期基因 115
细胞间黏附分子 1 164
细胞融合 251
细胞通讯 99
细胞因子 166
细胞应激 207
细胞增殖 114
细胞周期 115
细胞周期蛋白 115
细胞周期蛋白依赖性激酶 115
先天性缺陷 267
衔接蛋白 101
显性负效应 134
显性负效应突变 134
限制性核酸内切酶 19
限制性内切酶 233
限制性片段长度多态性 258

线粒体 DNA 24
线粒体 DNA 病 220
腺病毒 288
腺相关病毒 288
相连重叠群 251
小 G 蛋白 101
小干扰 RNA 19,86,278,285
心房颤动 184
心肌肥厚 182
心律失常 183
心钠素 176
心血管活性肽 175
锌指 83
新化学实体 299
信号识别颗粒 72
信号肽 71
信号肽酶 72
信号转导 99
信号转导分子 100
信号转导复合物 103
信号转导体 103
信号转导途径 99
信号转导子/转录活化子 110
信使 RNA 12
序列标签位点 90,253
序列图 90
选择素家族 164
选择性剪接 63
血管紧张素转换酶 176
血管内皮生长因子 118
血管重塑 176
血小板-内皮细胞黏附因子 1 164

Y

炎症介质 166
氧化应激 181,214
药物 296
药物蛋白质组学 303
药物发现 296
药物基因组学 302
药物敏感基因 294
药物设计 296
药物遗传学检测 268
一氧化氮 177
一致性序列 14
胰岛素 272
胰岛素依赖性糖尿病 203
遗传多态性 128
遗传密码 65
遗传图 89
遗传性着色性干皮病 49
乙型肝炎病毒 159
抑癌基因 140
抑素 119
抑制消减杂交 249
引发酶 29
引发体 29
应激 206
应激原 206
荧光原位杂交 251
由活性氧 214
有关线粒体 DNA 219
诱导 75
诱导性多能干细胞 126
原癌基因 141
原发性免疫缺陷病 195
阅读框移位 65

Z

杂化螺旋 60
杂化双链 227
载体 231
载脂蛋白 180
再生医学 126
早老素-1 222
造血干细胞 286
增强子 15,82
增色效应 227
增殖细胞核抗原 37
粘性末端 234
整合酶 288
整合素家族 164
脂多糖 150
脂氧素 173
脂质体 287
植入前遗传学诊断 267

质粒 25,233
致病岛 150
致病突变 128
中度重复序列 22
中性突变 129
终止密码子 65
肿瘤疫苗 281
重症联合免疫缺陷病 200
主控制基因 126
转分化 125
转基因动物 242
转基因技术 242
转录 12,56
转录辅激活因子 82
转录后基因沉默 86,244
转录激活因子 85
转录泡 60
转录图 90
转录抑制因子 85
转录因子 75,83
转录组 93
转录组学 88,303
转运 RNA 12
转运小体 189
着丝粒 12
阻遏 75
阻遏蛋白 76
组成性表达 75
组蛋白 10
组蛋白去乙酰化酶 82
组蛋白乙酰化酶 82
组合化学 301

英文索引

3′-UTR 133
α-synuclein 223
σ subunit 57

A

acquired immunodeficiency syndrome, AIDS 159
acute phase protein, APP 208
adapter proteins 101
adenovirus 288
adeno-associated virus, AAV 288
adult stem cell 125
ageing gene 217
allele specific oligonucleotide, ASO 262
alternative splicing 63
Alzheimer's disease, AD 220
aminoacyl site, A site 69
aminoacyl-tRNA synthetase 66
amplicon 250
amplified fragment length polymorphism, AFLP 260
amyloid precursor protein, App 221
anaphase promoting complex, APC 117
angiotensin converting enzyme, ACE 176
angiotensin Ⅱ, Ang Ⅱ 176
ankylosing spondylitis, AS 204
annealing 227
antisense RNA 19, 278
antisense technology 277
apolipoprotein, Apo 180
arrhythmia 183
asymmetric transcription 57
atherosclerosis, AS 177
atrial fibrillation, AF 184
atrial natriuretic peptide, ANP 176
attenuator 80

B

B cell receptor, BCR 195
bactericidal/permeability-increasing protein, BPI 276
base pair, bp 9
basic helix-loop-helix, bHLH 83
basic leucine zipper, bZIP 84
bioinformatics 92
biotechnology 226
blunt end 234
breast cancer gene 55

C

Ca^{2+}-dependent cell adhesion molecule family 164
calcitonin gene-related peptide, CGRP 168
calmodulin, CaM 108
cancer vaccines 281
cap binding proteins, CBPs 17
carboxyl terminal domain, CTD 58
cardiac hypertrophy 182
catabolite gene activation protein, CAP 78
causative mutation 128
CDK 39
cDNA library 231
cDNA microarray 249
cell communication 99
cell cycle checkpoint kinase 212
cell cycle 115
cell differentiation 120
cell division cycle genes, CDC genes 115
cell fusion 251
cell proliferation 114
cellular oncogene, c-onc 141
cellular stress 207
centromere 12
chalone 119
chemogenomics 304
chemokines 165
chemotaxis 165
cholera toxin, CT 158
chromatin 10

chromosome abnormalities 252
chromosome 10
cis-acting element 13
cis-regulation 13
cleavage and polyadenylation specificity factor, CPSF 62
cloning vector 233
closed transcription complex 58
coden 65
coding strand 57
codon 17
cohesive end 234
combinatorial chemistry 301
comparative genomics 88
complement receptor type 3 164
complementary DNA, cDNA 23
conditional gene knockout 244
congenital defects 267
consensus sequences 14
constitutive expression 75
contig 251
copy number variant, CNV 128
copy-number polymorphisms, CNP 128
core enzyme 57
co-activator 82
cyclin kinase inhibitor, CKI, or CDK inhibitory protein, CIP 117
cyclin 115
cyclin-dependent kinase, CDK 39,115
cyclooxygenase, Cox 168
cytokine 166

D

denaturing high performance liquid chromatography, DHPLC 263
dendritic cell, DC 293
deoxynucleotide triphosphate, dNTP 29
deoxyribonuclease, DNase 19
deoxyribonucleic acid, DNA 6
dideoxy-nucleoside triphosphate ddNTP 236
differential gene expression 75
disease related genes 246
disease susceptibility 128
DNA binding domain 83
DNA chip 92,262
DNA cloning 226
DNA damage 43
DNA fingerprinting 260
DNA ligase 29
DNA methylation 82
DNA polymerase 29
DNA recombination 226
DNA sequencing 236
DNA-dependent protein kinase, DNA-PK 51
DNA-dependent RNA polymerase 57
dominant negative effect 134
dominant-negative mutation 134
double helix 8
drug design 296
drug discovery 296
drug sensitive gene 294
drug 296
Duchenne muscular dystrophy DMD 253
dystrophin 253

E

efflux pump 153
embryo stem cell 242
embryonic stem cell 123,243
endonuclease 19
endothelin-1 176
endotoxin 150
enhancer 15
epidermal growth factor, EGF 118
epigenetic inheritance 82
EST 253
ex vivo 286
exon 12
exonuclease 19
exotoxin 150
expressed sequence tags, EST 90
expression vector 233

F

fluorescence in Situ hybridization FISH 251
frame shift 65

functional cloning 248
functional genomics 88
functional proteomics 95

G

G protein cycle 106
G protein 101
G protein-coupled receptor, GPCR 100
gain-of-function mutation 131
gaseous signal molecule 177
gene augmentation 285
gene cloning 233
gene correction 285
gene diagnosis 257
gene engineered subunit vaccine 279
gene engineered vaccine 279
gene expression regulation 74
gene expression 12,56
gene knockout 243
gene knock-in 243
gene location 251
gene manipulation 226
gene medicines 277
gene organization 12
gene replacement 285
gene targeting 243
gene therapy 284
gene transfection 236
gene transfer 285
gene vaccine 279
general transcription factors 85
genetic codes 65
genetic engineering 226
genetic map 89
genetic polymorphism 128
genetic program theory 217
genome 20
genomics 88
genotype 6,131
germline mutation 128
glutamic acid decarboxylase, GAD 204
glycomics 88
gross mutation 129
growth factors 118

H

haploid genotype 260
haplotype 131,260
heat shock element, HSE 210
heat shock factor, Hsf 210
heat shock protein, Hsp 71,133,209
helicase 29
helper virus 289
hematopoietic stem cell, HSC 286
Hepatitis B virus, HBV 159
herpes simplex virus, HSV 288
heteroduplex 227
high mobility group box 1, HMGB1 162
high throughput screening, HTS 301
highly repetitive sequences 22
histone acetyltransferase, HAT 82
histone deacetylase, HDAC 82
histone 10
holoenzyme 57
homologous recombination 51
housekeeping gene 75
human genome project, HGP 88
human immunodeficiency virus, HIV 159
human leukocyte antigen, HLA 195
hybrid helix 60
hyperchromic effect 227
hyperlipidemia 180
hypoxia-inducible factor, Hif-1 213

I

immunoglobulin superfamily 164
immunoglobulin, Ig 195
immuoreceptor tyrosine-based activation motif, ITAM 197
in silico cloning 253
in situ hybridization 251
in vivo 285
indirect diagnosis 258
induced pluripotent stem cell, iPS 126
inducible expression 242
induction 75

infectious diseases 149
inflammatory mediator 166
initiator codon 66
insulin 272
insulin-dependent diabetes mellitus, IDDM 203
integrase 288
integrated procedure for gene identification 249
integrin family 164
intracellular adhesion molecule-1 164
intron 13

J

Janus kinase, JAK 110

K

kilo-base pair, kbp 9
Kozac consensus sequence 68

L

lac operon 77
lateral inhibition 122
ligand-gated receptor channel 100
linkage analysis 252, 258
linkage disequilibrium 259
linkage map 89
lipopolysaccharide, LPS 150
liposome 287
lipoxin, LX 173
long interspersed repeated segments, LINES 23
long QT syndrome, LQTS 183
long terminal repeat, LTR 42
long termination repeat sequence, LTRs 288
longevity gene 217
loss-of-function mutation 131
low density lipoprotein, LDL 180
lymphocyte function-associated antigen-1, LFA-1 164

M

master control gene 126
mechanism-based drug design, MBDD 300
messenger RNA, mRNA 12
metabonomics 88, 304
microRNA, miRNA 19, 86
microsatellite 90
mitochondrial DNA disease 220
mitochondrial DNA, mtDNA 24
mitogen-activated protein kinase, MAPK 109
mitosis-promoting factor, MPF 117
moderately repetitive sequences 22
molecular chaperone 71
molecular cloning 226
monocistron 72
motif 83
mRNA differential display mRNA-DD 249
mRNA 小 RNA 19
mtDNA 219
multigene family 24
multiple cloning site MCS 233
mutation analysis 258

N

nerve growth factor, NGF 118
neurodegenerative disease 220
neutral mutation 129
new chemical entity, NCE 299
nitric oxide, NO 177
non-homologous end joining, NHEJ 51
Northern blot 240
nucleic acid vaccine 279
nucleic acid 6
nucleosome 10

O

Okazaki fragment 31
oligodeoxynucleotides, ODNs 277
oncogene 140
open reading frame, ORF 16, 65
open transcription complex 58
operator 14
operon 25
origin of replication 29
origin-recognition complex, ORC 36
oxidative stress 181, 214

P

Parkinson's disease, PD 220

pathogenicity islands, PAI 150
pathogenicity 150
pathogen-associated molecular patterns, PAMPs 162
pathological senility 216
PCR 261
peptide nucleic acid, PNA 278
peptidyl site, P site 69
peroxisome proliferator-activated receptor-γ, PPAR-γ 194
phage display 255
pharmacogenetics testing 268
pharmacogenomics 302
pharmacoproteomics 303
phenotype cloning 249
phenotype modulation 178
phosphoinositide 3-kinase related kinase, PIKK 211
phosphoinositide 3-kinase, PI3K 110
phospholipase C, PLC 100
photoreactivating enzyme 48
physical map 90
physiological senility 216
plasmid 25
platelet endothelial adhesion molecule-1 164
pluripotency 124
point mutation 129
poly(A) polymerase, PAP 62
poly(A)-binding protein, PABP 17
polyacrylamide gel electrophoresis PAGE 236
polycistron 25
polymerase chain reaction PCR 228
polyubiquitination 116
porin 153
positional cloning 251
post genomics 88
post-transcription gene silencing, PTGS 86
preimplantation genetic diagnosis, PGD 267
prenatal diagnosis 267
presenilin-1, PS-1 222
Pribnow box 14
primary immuno-deficiency diseases, PID 195
primase 29
primosome 29
probe 239
proliferation cell nuclear antigen, PCNA 37
promoter 14
prostaglandin, PG 168
protein kinase linked receptor 100
protein kinase 100
protein tyrosine kinase, PTK 100
proteome 93
proteomics 88
proteosomes 116
proto-oncogene 141

R

rational drug design 297
reactive oxygen species, ROS 44, 214
real-time PCR 239
receptor binding protein, RBP 155
receptor 100
recombinant staphylokinase, r-SAK 274
regenerative medicine 126
regulatory gene 76
regulatory region 12
release factor, RF 70
replication fork 30
replication licensing factor 39
replicon 30
representative difference analysis RDA 249
repressible gene 75
repression 75
repressor 76
restriction endonuclease 19
restriction fragment length polymorphism, RFLP 90, 258
retinoblastoma, RB 143
retrovirus 42, 288
reverse pharmacology 298
reverse transcriptase 23
ribonuclease, RNase 19
ribonucleic acid, RNA 6
ribosomal protein 18
ribosomal RNA, rRNA 12
ribosome cycle 69

ribosome 18
ribozyme 19
RNA dependent RNA polymerase RdRp 244
RNA editing 64
RNA interfering RNAi 244
RNA 244
RNase protection assay RPA 237
RNA-induced silencing complex, RISC 86
RNomics 19
RT-PCR 231

S

satellite DNA 22
second messenger 100
secondary immuno-deficiency diseases, SID 195
selectin family 164
semi-discontinuous replication 30
senescence, senility, ageing 216
sequence map 90
sequence tagged site, STS 90
sequence-tagged site STS 253
serial analysis of gene expression, SAGE 92,249
severe combined immunodeficiency diseases, SCID 200
Shine-Dalgarno sequence, SD sequence 69
short interspersed repeated segments, SINES 23
short tandem repeat mutation, STR mutation 129
short tandem repeat, STR 90
signal peptidase 72
signal peptide 71
signal recognition particle, SRP 72
signal transducer and activator of transcription, STAT 110
signal transducer 100
signal transduction pathway 99
signal transduction 99
signaling complex 103
signalsome 103
silencer 16
single copy sequences 22
single nucleotide polymorphism, SNP 90
single nucleotide polymorphisms, SNP 128
single-strand DNA binding protein 29
siRNA 244
site-directed mutagenesis 238
small G protein 101
small interference RNA, siRNA 86
small interfering RNA, siRNA 19
small non-messenger RNA, snmRNA 19
small nuclear ribonucleoprotein particle, snRNP 19
small nuclear RNA, snRNA 19
SNP 260
SNP-based haplotyping 260
somatic cell hybridization 251
somatic mutation 128
Southern blot 239
spatial specificity 75
special transcription factors 85
spliceosome 62
split gene 12
Src homology 2 domain 101
Src homology 3 domain 101
stable transfection 242
stage specificity 74
streptokinase, SK 274
stress 206
stressor 206
structural gene 12
structural proteomics 95
substance P, SP 168
superantigen 157
supercoils 9
suppressive substractive hybridization, SSH 249
system biology 96,305
systemic lupus erythematosus, SLE 205

T

T cell receptor, TCR 195
target 299
TATA box 15
telomerase 38
telomere 12,37
template strand 57
temporal specificity 74
terminator codon 65

Toll-like receptor 162
topoisomerase 29
totipotency 123
transcription activators 85
transcription bubble 60
transcription inhibitors 85
transcription map 90
transcription 12,56
transcriptome 93,303
transcriptomics 88
transdifferentiation 121,125
transfer RNA,tRNA 12
transgenic animal 242
transgenic technology 242
transient transfection 242
translation 12,56
transport vesicles 189
trans-acting factors 13
triplet code 17
triplex DNAs 9
trp operon 78
tumor suppressor gene 140

U

ubiquitin 116
untranslated region,UTR 17
urokinase-type plasminogen activator,u-PA 274

V

variation 128
vascular endothelial growth factor,VEGF 118
vascular remodeling 176
vasoactive peptide 175
vector 231
viropexis 155
virulence 149
virus oncogene,v-onc 141

W

Western blot 240

X

xeroderma pigmentosum,XP 49
X-linked agammaglobulinemia,XLA 198
X-linked hyper-IgM syndrome,XHIM 198

Y

yeast two-hybrid-system 255

Z

zinc finger 83